AD_____

Award Number DAMD17-98-1-8267

TITLE: Regulation of TGF-Beta Signal Transduction Pathways in Breast Cancer

PRINCIPAL INVESTIGATOR: Xuedong Liu, Ph.D.

CONTRACTING ORGANIZATION: Whitehead Institute for
Biomedical Research
Cambridge, Massachusetts 02142

REPORT DATE: June 1999

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	June 1999	Annual Summary (4 May 98 - 3 May 99)

4. TITLE AND SUBTITLE
Regulation of TGF-Beta Signal Transduction Pathways in Breast Cancer

5. FUNDING NUMBERS
DAMD17-98-1-8267

6. AUTHOR(S)
Xuedong Liu, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Whitehead Institute for Biomedical Research
Cambridge, Massachusetts 02142

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The goal of my studies under the BCRP fellowship is to understand the role of the Smad proteins and the TGF-ß signal ransduction pathways in breast cancer cell proliferation, tumor progression, and prognosis. In the past nine months, I have focused on identifying intracellular molecules and transcription factors that might be involved in the TGF-ß signal transduction pathway. Using both genetic and biochemical approaches, I have successfully isolated a few key molecules that are either mediator of the TGF-ß-induced transcriptional response or suppressors of the TGF-ß signal pathways. Future studies will be focused on understanding the detail mechanism of how the molecules interplay in mediating TGF-ß signaling and the significance of these interactions in tumor cell proliferation, progression, and prognosis.

14. SUBJECT TERMS
Breast Cancer, TGF-beta, Smad3, oncogene, PAI-1, promoter, retrovirus, cDNA library, expression cloning

15. NUMBER OF PAGES
181

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

__X__ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

__X__ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

__X__ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

PI - Signature Date

Table of Contents

The goal of my studies under the BCRP fellowship is to understand the role of the Smad proteins and the TGF-ß signal ransduction pathways in breast cancer cell proliferation, tumor progression, and prognosis. In the past nine months, I have focused on identifying intracellular molecules and transcription factors that might be involved in the TGF-ß signal transduction pathway. Using both genetic and biochemical approaches, I have successfully isolated a few key molecules that are either mediator of the TGF-ß-induced transcriptional response or suppressors of the TGF-ß signal pathways. Future studies will be focused on understanding the detail mechanism of how the molecules interplay in mediating TGF-ß signaling and the profound effect of these interactions in tumor cell proliferation, progression, and prognosis.

1. Genetic approach

A) Expression cloning of cDNAs encoding proteins or fragments of proteins from TGF-ß-resistant cancer cells that abolish the ability of TGF-ß to induce growth inhibition

TGF-ß exhibits strong antiproliferative effects on growth of many types of cells. Human fibrosacroma cell line HT1080 is growth insensitive to TGF-ß while exhibiting normal TGF-ß-induced transcriptional response, *i.e.,* induction of PAI-1. To understanding the reason why HT1080 cells are growth resistant to TGF-ß, my colleagues and I constructed a retroviral cDNA library from these cells and introduced the cDNA into a growth-sensitive mink lung epithelial cell line, L20. Upon selection with TGF-ß, we have obtained more than 20 clones that are growing in the presence of TGF-ß. I have isolated and sequenced some of the cDNAs in these cell lines. One of the frequently isolated genes is human activated N-ras oncogene with codon 61 Q to K mutation. I have reintroduced this mutant form of N-ras into the L20 cells, it recapitulated all the phenotypes I observed in the original clones. Incidentally, it has been shown previously that human HT1080 cells do harbor an activated N-ras mutation at codon 61. This result suggested that one of the reasons that HT1080 is growth resistant to TGF-ß is due to the mutation of N-ras. This result also validated our expression cloning strategy, and it indeed can clone the signaling molecules that overcome the growth inhibition response to TGF-ß. I have conducted similar screens using libraries constructed from other cell lines. I have cloned cDNA coding for the full length mdm2 gene, which has been reported recently by the David Beach group using a similar screen strategy. My next focus on this project is trying to elucidate what is the mechanism by which N-ras or mdm2 allows TGF-ß growth sensitive cells to become growth refractory.

B) Expression cloning of cDNAs encoding proteins participate in TGF-beta induced transcriptional response

TGF-ß is a potent inducer extracellular matrix production in a variety of cell types. One of the best studied examples is plasminogen activator inhibitor-1 (PAI-1). To identify proteins that participate in the TGF-ß-induced transcription of the PAI-1 gene, we established an expression cloning system involving an engineered cell line and a high titer retroviral cDNA library. TGF-ß-responsive BAH-gpt cells derived from HT1080 cells contain the bacterial guanosine phosphoribosyl transferase (gpt) gene under the control of 3TP, an artificial promoter derived from PAI-promoter region and 3XTRE element which has been shown previously mediating strong TGF-ß induced transcriptional response. These cells are deficient in hypoxanthine phosphoribosyl transferase (HPRT) and only grow in HAT medium and only in the presence of TGF-ß. We have introudced cDNA libraries into these cells and screened for cell clones that can proliferate in the absence of TGF-ß. One of the clones we have recovered and sequenced encodes human transcription

factor TFE3. We have subsequently shown that TFE3 can bind to the E-box region of the PAI-1 promoter and mediate strong transcriptional response to TGF-ß in collaboration with Smad3 protein. Thus, TFE3 is an important transcription factor in at least one TGF-ß-activated signal transduction pathway. This work has been recently published in Genes and Development 12:3084-3095,1998.

2. Biochemical Approach

Isolation and characterization a family protooncogene that specifically interact with Smad3 and regulate TGF-ß nuclear signaling

Previously, we have shown that TGF-ß induces rapid phosphorylation of Smad3 protein. Phosphorylation of Smad3 is required for TGF-ß induced transcriptional activation and growth inhibition. Mutation of ligand-dependent phosphorylation sites in Smad3 from serine to asparatic acid results in a constitutive active form of protein with regard to transcriptional activation. Using a GST-Smad3D recombinant protein as a bait in pull-down assays, we have identified a unique protein interact with Smad3D specifically in a ligand-dependent fashion. We have performed affinity chromatography and subsequently determined the probable identity of the protein by high sensitive Mass spectrometry sequencing methods. The protein we identify belongs to a family of proto-oncogene that regulates cell growth and proliferation. We have confirmed the strong ligand-dependent interaction between Smad3 and this family of proto-oncogenes. Overexpression of this family of protooncogenes can suppress the function of Smad3 and allow cells to become TGF-ß growth resistant. Ectopically overexpression of Smad3 can allow the growth-resistant cells to regain growth sensitivity of TGF-ß. Therefore, this family of proto-oncogenes can regulate cell proliferation through regulation of TGF-ß signaling pathways. Further studies are underway to elucidate the physiological significance of this interaction in cancer development.

The generous support from BCRP has made it possible for some of the work described here and is vital to continued success in understanding signaling mechanisms of TGF-ß and to achieve to the goals outlined in the fellowship application.

SUMMARY APPENDIX

Reportable Outcomes

1) Key Research Accomplishments

During the grant period I have made significant progress toward the goals I outlined in my grant application:

- I have constructed an effective expression cloning system based on retrovirus-mediated gene transfer;
- I have validated the system by successfully isolating several genes, such as mdm2 and N-ras that render sensitive parental cells to become resistant;
- I have made key contributions in identification and characterization of a transcription factor TFE3 and a family of proto-oncogenes that specifically interact with Smad 3 in regulating TGF-ß nuclear signaling.
- Finally, I am also involved in a research effort that leads to the elucidation of the mechanism of a viral oncogene gp55 activation of erythropoietin recepton.

2) Manuscripts, Abstracts, Presentations

Hua, X., **X. Liu,** D. Ansari, H. F.Lodish. Synergistic Cooperation of TFE3 and Smad Proteins in TGF-ß-Induced Transcription of the Plasminogen Activator Inhibitor-1 Gene. ***Genes and Development*** 12:3084-3095, 1998.

Wells, R. G., L. Giboa, Y. Sun, **X. Liu,** Y. Henis and H. F.Lodish. TGF-ß induces formation of a DTT-resistant type I/type II receptor complex in live cells. ***Journal of Biological Chemistry*** 274(9):5716-22,1999

Constantinescu, S. N., **X. Liu,** W. Beyer, A. Fallon, and H. F. Lodish. Point mutations in the membrane-spanning domain of the erythropoietin receptor alter susceptibility of Epo-dependent cells to viral oncogenesis. ***EMBO.*** In press (1999).

Liu, X, S.N. Constantinescu, J. Bogan, D. Hirsch and H. F. Lodish. Rapid generation of mammalian cells expressing multiple genes at predetermined levels Manuscript in preparation.

Sun*, Y. X. Liu*, E. Ng-Eaton, W.S. Lane, H.F. Lodish, and R. A.Weinberg. Nuclear oncoproteins interact with Smad3 regulates TGF-ß signaling . Manuscript in preparation. (* equally contributed)

N/A	Patents and licenses applied for and/or issued
N/A	Degrees obtained that are supported by this award
N/A	Development of cell lines, tissue or serum repositories
N/A	Informatics such as databases and animal models, etc
N/A	Funding applied for based on work supported by this award
N/A	Employment or research opportunities applied for and/or received based on experiences/training supported by this award

3) Copy of cited manuscripts - attached

Synergistic cooperation of TFE3 and Smad proteins in TGF-β-induced transcription of the plasminogen activator inhibitor-1 gene

Xianxin Hua,[1] Xuedong Liu,[1] Dominic O. Ansari,[1] and Harvey F. Lodish[1,2]

[1]The Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142 USA; [2]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 USA

GENES & DEVELOPMENT 12:3084–3095 © 1998 by Cold Spring Harbor Laboratory Press ISSN 0890-9369/98 $5.00; www.genesdev.org

Synergistic cooperation of TFE3 and Smad proteins in TGF-β-induced transcription of the plasminogen activator inhibitor-1 gene

Xianxin Hua,[1] Xuedong Liu,[1] Dominic O. Ansari,[1] and Harvey F. Lodish[1,2,3]

[1]The Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142 USA; [2]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 USA

Members of the TGF-β superfamily influence a broad range of biological activities including stimulation of wound healing and inhibition of cell growth. TGF-β signals through type I and II receptor serine/ threonine kinases and induces transcription of many genes including *plasminogen activator inhibitor-1* (*PAI-1*). To identify proteins that participate in TGF-β-induced gene expression, we developed a novel retrovirus-mediated expression cloning strategy; and using this approach, we established that transcription factor μE3 (TFE3) is involved in TGF-β-induced activation of the *PAI-1* promoter. We showed that TFE3 binds to an E-box sequence in PE2, a 56-bp promoter fragment of the *PAI-1* promoter, and that mutation of this sequence abolishes both TFE3 binding as well as TGF-β-dependent activation. TFE3 and Smad3 synergistically activate the PE2 promoter and phosphorylated Smad3 and Smad4 bind to a sequence adjacent to the TFE3-binding site in this promoter. Binding of both TFE3 and the Smad proteins to their cognate sequences is indispensable for TGF-β-inducible activation of the PE2 promoter. Hence, TFE3 is an important transcription factor in at least one TGF-β-activated signal transduction pathway.

[*Key Words*: TFE3; Smads; TGF-β; E box; *PAI-1*]

Received June 25, 1998; revised version accepted August 5, 1998.

TGF-β has a diverse range of biological activities including inhibition of cell growth, induction of cell differentiation, inhibition of the immune response, and production of extracellular matrix proteins. TGF-β rapidly induces transcription of extracellular matrix proteins such as plasminogen activator inhibitor-1 (PAI-1) and cell-cycle inhibitors such as $p15^{INK4B}$ and $p21^{WAF1/CIP}$ (Attisano et al. 1994; Hannon and Beach 1994; Datto et al. 1995; Li et al. 1995). TGF-β signals through the sequential activation of two homodimeric cell-surface receptors, termed type I and type II (TβRI and TβRII), both of which are serine–threonine protein kinases (Franzen et al. 1993; Lin and Lodish 1993; Wrana et al. 1994; Luo and Lodish 1996; Weis-Garcia and Massagué 1996). Ligand-activated TβRI phosphorylates conserved serines at the carboxyl termini of either Smad2 or Smad3, which promote their binding to one or more molecules of Smad4, a common partner for all phosphorylated Smads involved in signaling by both TGF-β and bone morphogenetic proteins (Heldin et al. 1997; Massagué et al. 1997; Attisano and Wrana 1998). Smad complexes then enter the nucleus and activate transcription of a variety of genes.

Ectopic expression of Smad2 and Smad4 activates transcription of a reporter gene driven by the artificial TGF-β responsive promoter 3TP, and addition of TGF-β further stimulates its expression (Lagna et al. 1996; Macias-Silva et al. 1996). Similarly, overexpression of Smad3 and Smad4 activates transcription from the TGF-β inducible *PAI-1* promoter (Zhang et al. 1996; X. Liu et al. 1997). Recent evidence indicates that Smad3 and Smad4 can directly bind to specific DNA sequences in either artificial or natural TGF-β-inducible promoters and thus activate gene transcription (Yingling et al. 1997; Dennler et al. 1998; Vindevoghel et al. 1998; Zawel et al. 1998). However, it is not yet clear whether binding only of Smad proteins to DNA is sufficient to confer maximal TGF-β-induced transcription. Given the diversity of genes that are activated by TGF-β, it seems likely that other transcription factors, some possibly expressed only in certain kinds of cells, partner with phosphorylated Smads to induce the expression of different genes. In *Xenopus*, activin induces phosphorylation of Smad2, which then forms a complex with Smad4 and the tran-

[3]Corresponding author.
E-MAIL lodish@wi.mit.edu; FAX (617) 258-6768.

scription factor FAST-1. This complex then binds to the promoter of the developmentally regulated gene *Mix2* and induces its transcription (X. Chen et al. 1996, 1997; F. Liu et al. 1997).

PAI-1 is a component of the extracellular matrix and also plays an important role in regulating blood coagulation. Addition of TGF-β to cultured human Hep G2 hepatoma cells dramatically induces *PAI-1* gene expression (Westerhausen et al. 1991). Coexpression of Smad2 or Smad3 with Smad4 also increases the expression of reporter genes driven by the *PAI-1* promoter (Y. Chen et al. 1996; Lagna et al. 1996; Zhang et al. 1996), and mutation of Smad4 leads to loss of the TGF-β response (F. Liu et al. 1997; Zawel et al. 1998). Thus, Smad proteins are involved in TGF-β-induced transcription of the *PAI-1* gene. AP-1 binding sites as well as an E-box sequence have been implicated in TGF-β-induced transcription of the *PAI-1* gene (Keeton et al. 1991; Riccio et al. 1992). However, little is known about how transcription factors in concert with Smad proteins stimulate TGF-β-dependent *PAI-1* transcription.

We developed a novel expression cloning system involving an engineered TGF-β responsive cell line and a retroviral cDNA library. Using this system, we cloned a transcription factor, TFE3, which when ectopically expressed activates TGF-β-induced expression of the *PAI-1* gene. We identified two E-box DNA sequences in the natural *PAI-1* promoter that specifically bind to TFE3. Ectopic expression of TFE3 increases TGF-β-dependent expression of a reporter gene driven by a natural *PAI-1* promoter fivefold, whereas mutation of the E-box sequence in a fragment of the *PAI-1* promoter completely abolishes TGF-β-inducible transcription. Moreover, we showed that TFE3 and Smad3 and Smad4 synergize in enhancing TGF-β-dependent transcription from this minimal 56-bp promoter fragment, a synergy strictly dependent on the phosphorylation of serine residues at the carboxyl terminus of Smad3. Furthermore, we show that within a 36-bp *PAI-1* promoter, a complex of Smad3 and Smad4 bind to a sequence adjacent to the TFE3-binding site. Binding of this DNA by a Smad3–Smad4 complex requires phosphorylation of the carboxyl terminus of Smad3, and binding of both TFE3 and Smad proteins to this promoter is essential for TGF-β-inducible transcription. Together, our data show that TFE3 and Smad proteins synergistically cooperate in transcription of at least one TGF-β-inducible gene.

Results

Isolation of cell clones with constitutive TGF-β signaling on infection with a retroviral cDNA library

To identify proteins that participate in the TGF-β-induced transcription of the *PAI-1* gene, we established an expression cloning system involving an engineered cell line and a high titer retroviral cDNA library. Our expression cloning strategy was based on the finding that overexpression of either Smad2 or Smad3 alone or together with Smad4 induces TGF-β responses in the absence of ligand (Y. Chen et al. 1996; Lagna et al. 1996; Zhang et al. 1996; X. Liu et al. 1997); thus overexpression of other proteins in the TGF-β-signaling pathway might also induce transcription of genes normally activated by TGF-β. To infect the engineered TGF-β-responsive BAH-gpt (guanosine phosphoribosyl transferase) cells (Hocevar and Howe 1996) with a retroviral cDNA library, we stably introduced into the cells the cDNA encoding the receptor for the ecotropic murine Moloney retrovirus (Baker et al. 1992). Growth of the resulting BAH-ER3 cells, like parental BAH-gpt cells, was not inhibited by TGF-β but transcription of the *PAI-1* gene was fully inducible by TGF-β. These cells are deficient in hypoxanthine phosphoribosyl transferase (HPRT), but contain the bacterial *gpt* gene under the control of 3TP, an artificial TGF-β inducible promoter (Wrana et al. 1992; Hocevar and Howe 1996) (Fig. 1). Thus, BAH-ER3 cells grow in HAT medium only in the presence of TGF-β. In contrast, addition of 6-thioguanine (6-TG) to normal medium kills the cells, but only in the presence of TGF-β, as *gpt* converts 6-TG to a toxic product (Fig. 1).

Two million BAH-ER3 cells were infected with a retroviral cDNA library prepared from HPRT-deficient HT1080 cells, and grown in HAT medium in the absence of TGF-β. After 2 weeks, we isolated 12 HAT-resistant clones. We reasoned that if overexpression of a particular cDNA caused the HAT-resistant phenotype, then we could rescue the cDNA sequence by packaging of the retroviral RNA by the Gag, Pol, and Env proteins derived from superinfecting wild-type Moloney retroviruses (Rasheed 1995). The rescued retrovirus transducing the cDNA should allow transfer of the HAT-resistant phenotype to fresh BAH-ER3 cells. Retroviruses produced by 3 of the 12 lines, including HATR4 and HATR7 cells, caused BAH-ER3 cells to acquire the ability to grow in

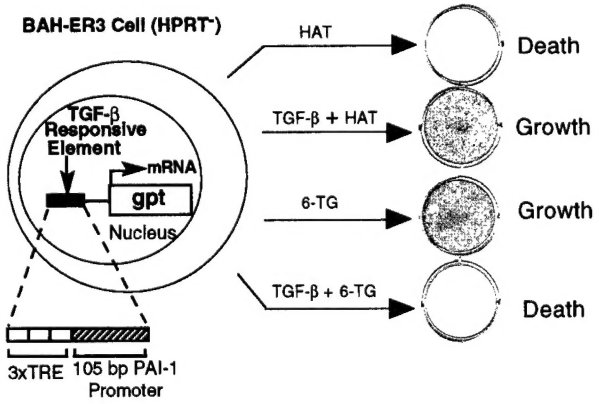

Figure 1. TGF-β-regulated growth of BAH-ER3 cells in the presence of drug selection. On day 0, BAH-ER3 cells were seeded at a density of 5×10^4 cells/well in a six-well plate in DMEM containing 10% fetal calf serum, 100 U/ml penicillin and 100 μg/ml streptomycin. On day 1, the cells were switched to medium with 1× HAT or 6-TG (30 μg/ml) with or without 200 pM TGF-β as indicated. On day 9, the growing cells were stained with crystal violet. (TRE) Phorbol ester TPA response element; (gpt) guanosine phosphoribosyl transferase.

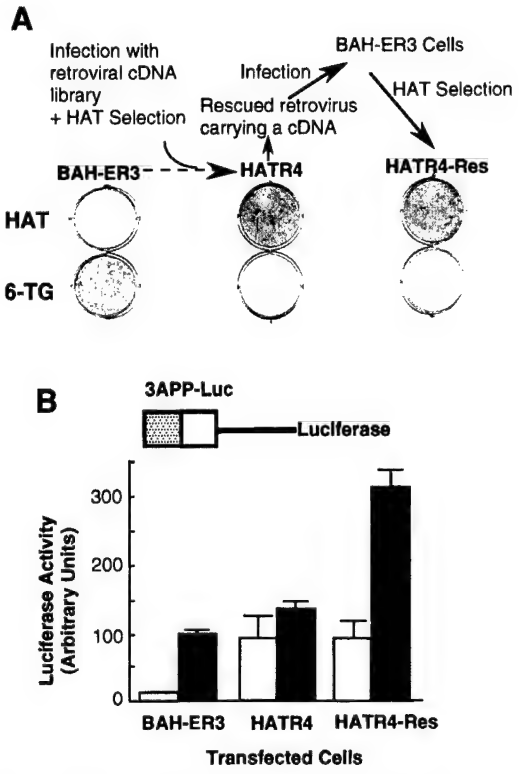

Figure 2. Isolation of a HAT-resistant cell clone that activates the TGF-β-inducible 3TP promoter in the absence of TGF-β. (*A*) After infection of BAH-ER3 cells with a retroviral cDNA library, the cells were grown in HAT medium for 2 weeks. A HAT-resistant clone, HATR4, was isolated, and then infected with wild-type Moloney retrovirus; the supernatant containing the rescued retrovirus was used to infect normal BAH-ER3 cells as described in Materials and Methods. Infected cells were also subjected to HAT selection and the resulting HAT-resistant cells, HATR4–Res cells, were isolated. BAH-ER3, HATR4, and HATR4–Res cells were plated at 5×10^4 cells/well in a 6-well plate; HAT medium or 6-TG medium was added as indicated, and cells were incubated for 10 days in the absence of TGF-β. Growing cells were stained with crystal violet. (*B*) On day 0, 10^5 cells were plated in each well of a 12-well plate. On day 2, the cells in each well were transfected with 2 μg of 3APP–Luc DNA and 0.5 μg of pSVβ. After overnight culture the cells were switched to serum-free medium with (■) or without (□) 200 pM TGF-β as indicated, and then incubated for 20 hr before being harvested for luciferase and β-galactosidase assays as described in Materials and Methods. Luciferase activities, plotted in arbitrary units, have been normalized to β-galactosidase activity. Each value represents an average of duplicate samples, and the error bar denotes the standard deviation of the duplicates.

HAT medium in the absence of TGF-β and to be killed in normal medium in the presence of 6-TG (Fig. 2A).

The HAT-resistant HATR4-Res cells were derived from the BAH-ER3 cells infected with the retrovirus rescued from HATR4 cells (Fig. 2A). If constitutive expression of the *gpt* gene is caused by constitutive activation of the TGF-β-inducible promoter upstream of the *gpt* gene, then we expect that a similar TGF-β-inducible promoter in a luciferase reporter construct should drive expression of luciferase even in the absence of TGF-β. To test this hypothesis, we transfected a TGF-β-inducible luciferase reporter construct, 3APP–Luc, into parental BAH-ER3 cells, HATR4 cells, and the rescued HATR4–Res cells (Fig. 2B). In BAH-ER3 cells, expression of the TGF-β-inducible 3APP–Luc is low in the absence of TGF-β and induced sevenfold by TGF-β. In contrast, in the absence of TGF-β expression of 3APP–Luc is much higher in both HATR4 and HATR4–Res cells than that in BAH-ER3 cells. Addition of TGF-β has little effect on expression of the reporter gene in HATR4 cells. In HATR4–Res cells, which may contain multiple copies of the retroviral genome, TGF-β stimulates reporter gene expression threefold (Fig. 2B). These results suggest that a retrovirus-introduced cDNA is responsible for HAT resistance as well as constitutive expression of the normally TGF-β-inducible reporter gene.

Cloning of TFE3 from the cell clone with constitutive TGF-β signaling

To clone the cDNA responsible for the HAT-resistant phenotype of HATR4 cells by use of PCR, we amplified genomic DNA with a pair of oligonucleotides flanking the multiple cloning site in the retroviral vector. A single 2.7-kb DNA fragment was amplified from the genomic DNA of HATR4 cells, but not from control BAH-ER3 cells (data not shown). Sequencing of this DNA fragment indicates that it encodes the full-length transcription factor μE3 (TFE3), a ubiquitously expressed basic helix–loop–helix transcription factor originally isolated as a factor binding to the E-box sequence (CACGTG) in the enhancer of an immunoglobulin gene (Beckmann et al. 1990; Zhao et al. 1993). A 1.9-kb cDNA encoding the full-length TFE3 was also cloned from HATR7 cells (data not shown).

TFE3 enhances TGF-β-dependent activation of the PAI-1 *promoter*

To determine whether TFE3 activates the expression of the luciferase reporter gene driven by the natural *PAI-1* promoter, which is well induced by TGF-β (Keeton et al. 1991; Westerhausen et al. 1991; Riccio et al. 1992), we transfected PAI-Luc into BAH-ER3 cells. Coexpression of TFE3 enhanced TGF-β-independent expression of PAI–Luc less than twofold (Fig. 3B). Importantly, cotransfection of TFE3 enhanced *PAI-1* promoter activity fivefold in the presence of TGF-β. This suggests that TFE3 is involved in TGF-β-induced transcription of the *PAI-1* gene.

To identify the minimal element(s) in the *PAI-1* promoter that are responsive to both TFE3 and TGF-β, we tested the activity of three fragments of the full-length *PAI-1* promoter. Figure 4B shows that activity of fragment PF1 (bases −794 to −532) of the *PAI-1* promoter, which contains two perfect TFE3-binding E-box sequences (CACGTG), is stimulated fivefold by TGF-β. Notably, coexpression of TFE3 enhances PF-1 promoter

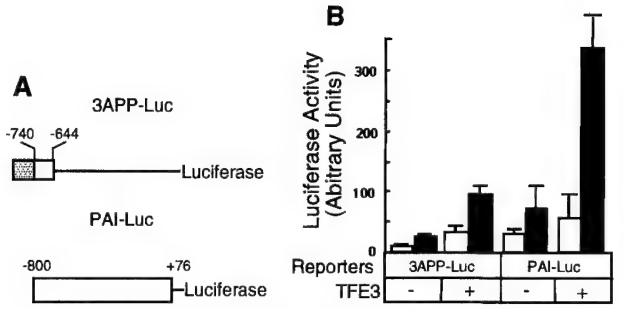

Figure 3. TFE3 activates the *PAI-1* promoter in a TGF-β-dependent fashion. (*A*) A diagram of the luciferase reporter genes driven by various promoters. The three tandem repeats of a 7-bp AP1-binding site [(stippled box) TGA(G/C)TCA] separated by an *Xba*I site were inserted upstream of −740 to −644 fragment of the *PAI-1* promoter (open box) in the 3APP–Luc construct. Hence, the short AP1-binding sequence was used to replace the 32-bp fragment containing an AP1-binding sequence in the 3TP promoter (Wrana et al. 1992) to eliminate the potential influence of sequences other than the AP1-binding site. (*B*) Activation of PAI–Luc expression by TFE3 is dependent on TGF-β. BAH-ER3 cells were transfected as described in Fig. 2B. Each well received 1.5 μg of reporter gene, 0.5 μg of pSVβ, and also 1.0 μg of plasmid encoding TFE3 as indicated. The total amount of DNA per well was adjusted to 3.0 μg. Transfected cells were treated with (■) or without (□) TGF-β, and processed for both luciferase and β-galactosidase assays as described in Fig. 2B.

activity dramatically in the presence of TGF-β but only slightly in the absence of TGF-β. In contrast, segments PF2 and PF3, containing bases −552 to −194 and −214 to +29, respectively, are not responsive to TGF-β and are unaffected by TFE3 overexpression (Fig. 4B). Subdivision of the PF1 promoter into smaller pieces showed that at least two subfragments, PE1 and PE2, each of which contains one E box, are responsive to TGF-β; this experiment (Fig. 4C) was done by use of Hep G2 cells because we found that the expression of these luciferase reporter genes were more regulatable by TGF-β in this cell line. The higher level of expression of PF1–Luc compared with PE1–Luc and PE2–Luc is probably the result of the effect of tandem repeats of PE1 and PE2 in the PF1 promoter fragment.

E box, the TFE3-binding sequence, is essential for TGF-β-induced activation of the PE2 promoter

Each of the two TGF-β-responsive elements, PE1 and PE2, contains a perfect E-box sequence. We tested the importance of the E box in the PE2 promoter by transfecting luciferase reporter genes driven by the wild-type PE2 promoter (PE2–Luc) or a promoter with the mutant E box (PmE2-Luc, CACGTG → acCGac) (Fig. 5A). All activity of the PE2 promoter was dependent on the presence of a functional E box, because the mutant was inactive. In contrast, activity of the wild-type PE2 promoter was stimulated fourfold by TFE3 in the presence of TGF-β (Fig. 5B). This result is consistent with a previous report showing that the E-box sequence in the PE2 fragment of the *PAI-1* promoter is critical for TGF-β-induced transcription of the *PAI-1* gene (Riccio et al. 1992).

The gel-shift assay in Figure 5C shows that TFE3 synthesized in vitro binds to a ^{32}P-labeled PE2 DNA probe (Fig. 5, lane 2). Binding was competed completely by unlabeled wild-type PE2 oligonucleotides (Fig. 5, lane 3) but not by oligonucleotides bearing a scrambled mutation in the E-box sequence (CACGTG → acCGac; Fig. 5, lane 4). Together, these data suggest that the E-box sequence in the PE2 promoter is essential for the binding of TFE3 to the promoter as well as for TFE3- and TGF-β-dependent activation of the promoter.

TFE3 and Smad3–Smad4 synergize in TGF-β-dependent transcription

The data in Figure 6A show that TFE3 synergizes with Smad3 in enhancing TGF-β-dependent activation of the PE2 *PAI-1* promoter. We transfected Hep G2 cells with the PE2-Luc reporter and various Smad constructs.

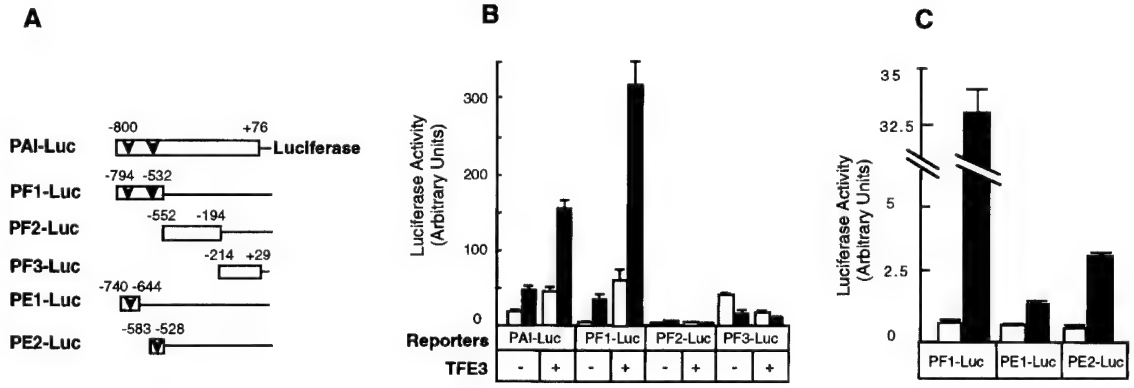

Figure 4. Identification of small subfragments, PE1 and PE2, of the *PAI-1* promoter that are activated by TFE3 and TGF-β. (*A*) A diagram of the reporter constructs. (Open bars) *PAI-1* promoter; (arrowhead) E-box sequence (CACGTG). BAH-ER3 cells (*B*) and Hep G2 cells (*C*) were transfected and the luciferase and β-galactosidase assays were carried out as detailed in Fig. 3B. (□) Without TGF-β; (■) with TGF-β.

Transfection of either Smad3 or Smad4 or both together had little effect on the PE2 promoter activity, either in the absence or presence of TGF-β. Transfection of a small amount of TFE3 plasmid DNA alone slightly stimulated the PE2 promoter activity in the presence of TGF-β. Importantly, cotransfection of TFE3 and Smad3, or TFE3, Smad3, and Smad4 together, markedly stimulated the PE2 promoter activity in the presence of TGF-β. In contrast, only a slight stimulation was observed in the absence of TGF-β. The reporter construct PmE2–Luc, containing a mutant E box, was inactive even after cotransfection of TFE3, Smad3, and Smad4 and stimulation with TGF-β.

TGF-β induces phosphorylation of the serine residues at the carboxyl terminus of Smad3. Phosphorylation is essential for signaling because overexpression of the mutant Smad3A, in which the three carboxy-terminal serines are changed to alanines, blocks the ability of TGF-β to inhibit cell division and stimulate the *PAI-1* promoter (X. Liu et al. 1997). Consistent with these observations, Smad3A had little effect on the PE2 promoter activity, either in the absence or presence of TFE3 or TGF-β (Fig. 6B). Taken together, these results show a functional synergy between TFE3 and phosphorylated Smad3 in activation of the PE2 promoter.

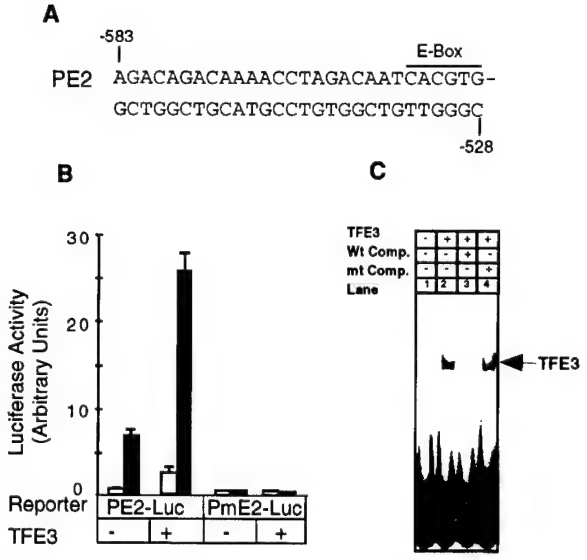

Figure 5. The E-box sequence is essential for TFE3-mediated and TGF-β-dependent activation of the PE2 fragment of the *PAI-1* promoter. (*A*) The sequence of the PE2 promoter. (*B*) Hep G2 cells were transfected with 1 μg of PE2–Luc or PmE2–Luc DNA (CACGTG → acCGac), 1 μg of pSVβ, and also 1.0 μg of plasmid encoding TFE3 as indicated. Transfected cells were treated with (■) or without (□) TGF-β, processed, and assayed as described in Fig. 3B. (*C*) TFE3 was synthesized in vitro from pET–TFE3 by the TNT T7 Coupled Reticulocyte Lysate System (Promega). Gel-shift reactions contained 3 μl of the in vitro TFE3 translation reaction and 1 μl of (4×10^3 cpm) of ^{32}P-labeled PE2 DNA probe. Reactions in lanes *3* and *4* contained a 50-fold excess of wild type or mutant PE2 oligonucleotides, respectively.

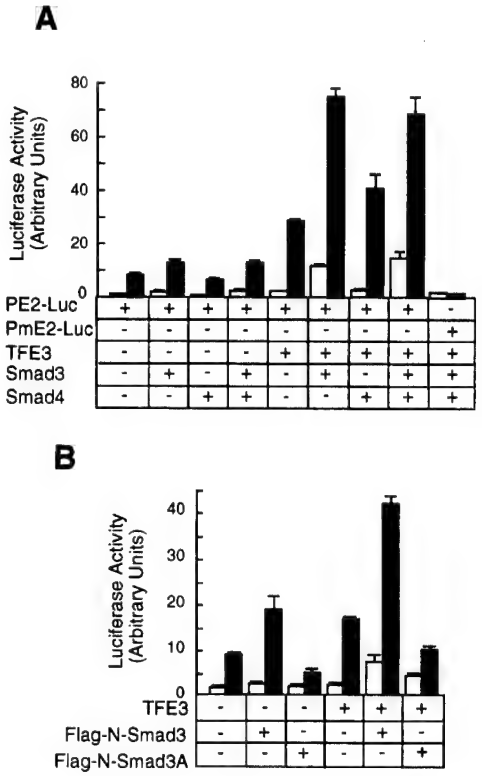

Figure 6. Synergy between TFE3 and Smad3 in the activation of the PE2 promoter. (*A*) Hep G2 cells were transfected as described in the legend to Fig. 5A. The following plasmids were used in transfection as indicated: 0.5 μg of plasmid encoding TFE3, 1 μg of plasmid encoding Smad3, and 1 μg of plasmid encoding Smad4; every well received 1 μg of PE2–Luc and 0.2 μg of pCMV–β-gal. The total amount of DNA was adjusted to 3.7 μg per well with pcDNA3. (*B*) Hep G2 cells were transfected with the following plasmids: 0.5 μg of TFE3, 1 μg of Flag-N–Smad3 or Flag-N–Smad3A; every well received 1 μg of PE2–Luc and 0.2 μg of pCMV-β. The total amount of DNA per well was adjusted to 3.7 μg with a control plasmid, pEXL–GFP. The cells were transfected, treated with (■) or without (□) TGF-β, and assayed as described for panel *A*.

Smad 4 and phosphorylated but not unphosphorylated Smad3 together bind to the PE2.1 element of the PAI-1 *promoter*

The PE2 fragment of the *PAI-1* promoter, bases −583 to −528, contains a perfect E box at −561 to −556. As detailed below, we surmised that a Smad3–Smad4 complex binds to nucleotides within −583 to −528; thus, we tested the PE2.1 probe, containing two tandem segments of DNA spanning bases −586 to −551 of the *PAI-1* promoter. The gel-shift experiment in Figure 7A shows that a complex of Smad4 and phosphorylated Smad3 binds to this DNA fragment. In this study we transfected Bosc23 cells with plasmids encoding Smad4 and/or Flag-tagged Smad3, together with the constitutively active type I TGF-β receptor, TβRI–T204D. Lysates from transfected cells were then incubated with the ^{32}P-labeled PE2.1 probe and analyzed on a native polyacrylamide gel (Fig. 7).

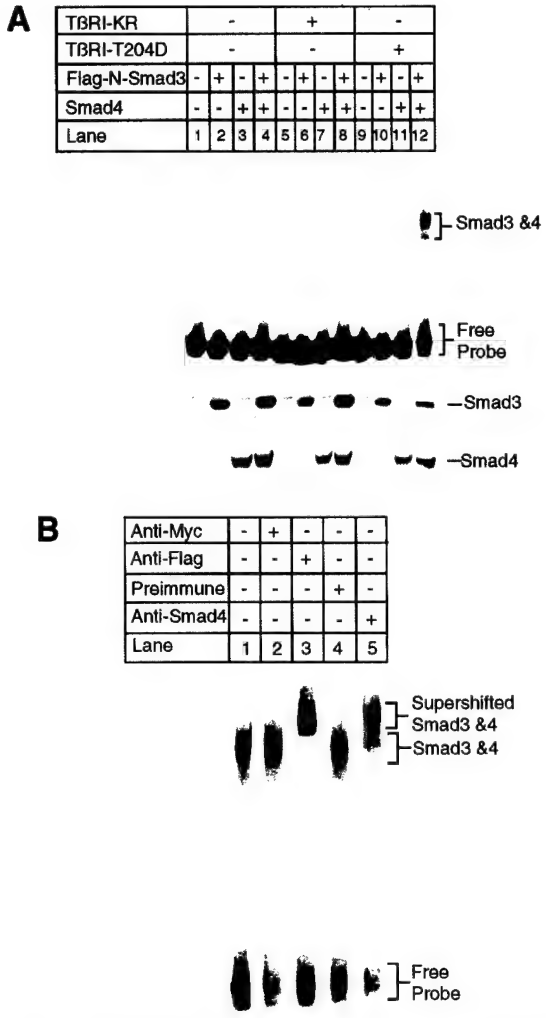

Figure 7. Smad3 and Smad4 together bind to the PE2.1 element of the *PAI-1* promoter. (*A*) Bosc23 cells were transfected as described in Materials and Methods; cells in each dish were transfected with 2 μg of plasmid encoding Flag-N-Smad3 or Smad4 (pEXL–Smad4), and 1 μg of plasmid encoding TβRI-KR (pCMV5–TβRI–KR) or TβRI-T204D (pCMV5-TβRI-T204D) as indicated. The total amount of DNA for each dish was adjusted to 5 μg with pEXL–GFP. The gel-shift assay at *top* was carried out with the ^{32}P-labeled probe and 1 μl of cell lysate as described in Materials and Methods, and exposed to a Fuji PhosphorImager plate. The minor lower band in lane *12* at *top* probably represents Smad3 and Smad4 protein binding to only one of the two tandem repeats of the PE2.1 element. (*Middle, bottom*) Immunoblots with 5 μl (150 μg of proteins) of cell lysates in each lane that were blotted with an anti-Flag M2 antibody, recognizing the Flag epitope-tagged Smad3, and with an anti-Smad4 antibody, respectively, as described in Materials and Methods. As indicated, the levels of expression of Smad4 were the same in all cases (lanes *3,4,7,8,11,12*) as were those of the Flag-tagged Smad3 (lanes *2,4,6,8,10,12*). (*B*) Cell lysates containing both Smad4 and the Flag-tagged Smad3 were incubated with ^{32}P-labeled PE2.1 DNA, followed by addition of 1 μl of an irrelevant control antibody or preimmune serum (lanes *2,4*) or the anti-Flag antibody (anti-Smad3) (lane *3*) and the anti-Smad4 antibody (lane *5*), followed by gel electrophoresis as described in Materials and Methods.

Only lysates from cells expressing Smad3, Smad 4, and the active TβRI–T204D bound this PE2.1 element (Fig. 7A, lane 12); lysates from cells transfected with either Smad3 or Smad4 failed to bind (Fig. 7A, lanes 10,11), indicating that a complex of Smad 3 and Smad4 is binding to this probe. This gel-shifted complex can be supershifted by either an anti-Flag antibody, recognizing the epitope-tagged Smad3, or by an anti-Smad4 antibody (Fig. 7B, lanes 3,5), but not by control antibodies (Fig. 7B, lanes 2,4), confirming the presence of both Smad3 and Smad4 in the complex. Figure 7A, lanes 4–8 provide additional controls, showing that cells transfected with a kinase-deficient type I receptor fail to generate a functional DNA-binding complex. Furthermore, cotransfection of cells with TβRI–T204D, Smad4, and the mutant Smad3A did not yield a complex capable of binding the PE2.1 probe (Fig. 8, lane 6). Dennler et al. (1998) reported that GST fusion proteins of both full-length Smad4 and the amino-terminal half of Smad3 independently and directly bind to multiple CAGA sequences derived from the *PAI-1* promoter. There is one CAGA sequence in the PE2.1 fragment. In contrast, our experiments show that binding of Smad3 and Smad4 to the 36-bp PE2.1 promoter fragment depends on the presence of the consti-

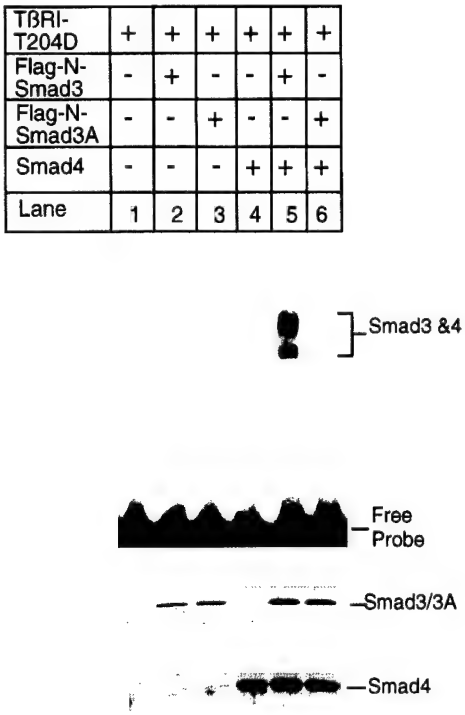

Figure 8. Phosphorylation of Smad3 at the carboxyl terminus is essential for binding of Smad3 and Smad4 to the PE2.1 element of the *PAI-1* promoter. Bosc23 cells were transfected with the constitutively active TβRI–T204D, Smad4, and wild-type Smad3 or mutant Smad3A as described in Fig. 7. (*Top*) Results of a gel shift assay. (*Middle, bottom*) Immunoblots with 5 μl (150 μg of protein) of cell lysates in each lane that were blotted with an anti-Flag M2 antibody, recognizing the Flag epitope-tagged Smad3, and with an anti-Smad4 antibody, respectively, as described in the legend to Fig 7A.

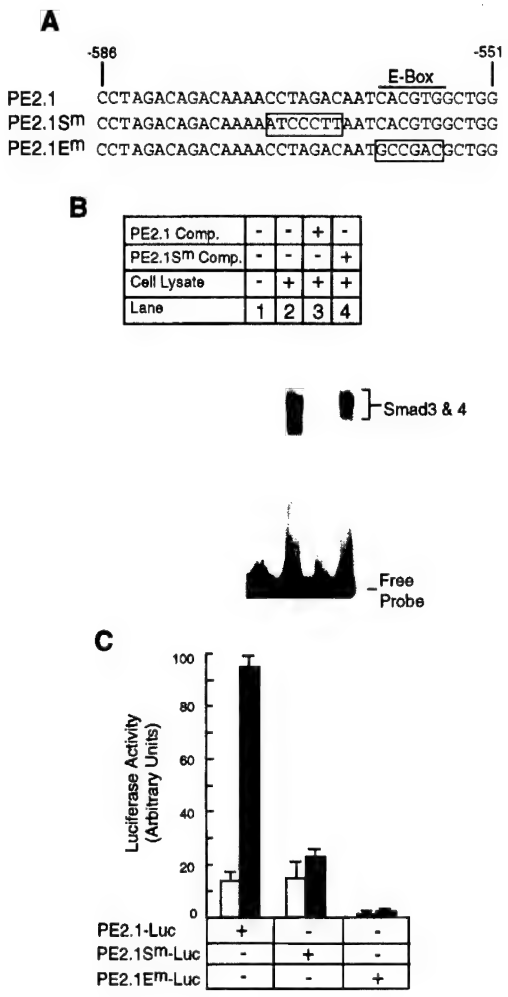

Figure 9. Both the Smad3–Smad4 binding site and the TFE3 binding site are essential for TGF-β-induced activation of the 36 nucleotide PE2.1 element of the *PAI-1* promoter. (*A*) The sequences of one of the two identical repeats of the wild-type and two mutant PE2.1 elements. (*B*) Lysate from cells overexpressing TβRI-T204D, Smad3, and Smad4 as described in Fig. 7 were incubated with radiolabeled PE2.1 DNA alone (lane *2*) or in the presence of unlabeled PE2.1 (lane *3*) or mutant PE2.1S^m oligonucleotides (lane *4*), followed by gel electrophoresis. Both the radiolabeled DNA and the competing DNA contained two tandem repeats of the corresponding sequence. (*C*) Luciferase reporter genes driven by two tandem repeats of either the PE2.1, PE2.1S^m, or PE2.1 E^m elements were transfected into Hep G2 cells, and luciferase assays were carried out as described in Materials and Methods. (■) With TGF-β; (□) without TGF-β.

tutively active type I receptor, and that the complex contains Smad4 and phosphorylated Smad3.

Binding of TFE3 and Smad3–Smad4 to adjacent sites in the 36 bp of the PE2.1 element induces TGF-β-dependent transcription

Figure 9 shows that TFE3 and the complex of Smad4 and phosphorylated Smad3 bind to adjacent sequence of the PE2.1 promoter, and that both binding sites are essential for TGF-β-induced promoter activity. This experiment compares the activity of wild-type PE2.1 DNA (–586 to –551) with that of its two mutant versions, namely a mutation in the potential Smad3 and Smad4 binding site, PE2.1S^m, and a mutation in the E box, PE2.1E^m (Fig. 9A). Lysates from cells overexpressing Smad4 and phosphorylated Smad3 form a gel-shifted band with radiolabeled PE2.1 DNA (Fig. 9B, lane 2) that is blocked by an excess of unlabeled PE2.1 DNA (Fig. 9B, lane 3), but not by unlabeled PE2.1S^m DNA (Fig. 9B, lane 4). This indicates that at least part of the sequence 5′-CCTAGAC-3′, located 3 nucleotides in front of the E-box sequence, is bound by the complex of Smad4 and activated Smad3.

Figure 9C shows that two tandem repeats of the PE2.1 sequence, inserted upstream of a luciferase reporter gene, forms a functional TGF-β-inducible promoter. Mutation of the E boxes in this sequence inactivated promoter activity, both in the presence and absence of TGF-β. In contrast, mutation of the Smad binding site, PE2.1S^m, had no effect on basal (TGF-β-independent) promoter activity, but abolished the ability of TGF-β to stimulate transcription.

Together, these experiments indicate that both the Smad3–Smad4-binding site and the TFE3-binding site in PE2.1, the 36-bp TGF-β-inducible *PAI-1* promoter, are essential for TGF-β-induced gene activation, suggesting that TFE3 synergizes with the Smad3–Smad4 protein complex in TGF-β signaling by binding to adjacent sites in the promoter. The gel-shift experiment in Figure 10

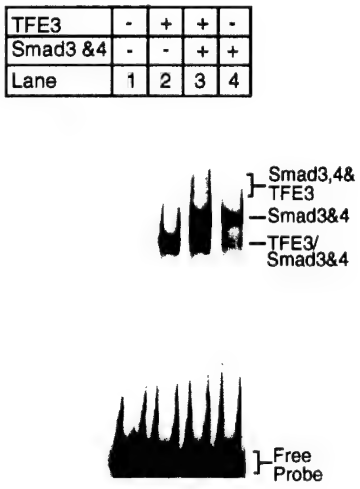

Figure 10. Both TFE3 and Smad3–Smad4 bind to the PE2.1 promoter fragment of the *PAI-1* promoter. TFE3 was synthesized in vitro with a Promega TNT kit as instructed by the manufacturer. In vitro synthesized TFE3 (2 μl; lanes *2,3*), and/or lysates from cells overexpressing TβRI-T204D, Smad3, and Smad4 (1 μl; lanes *3,4*) were incubated with radiolabeled PE2.1 DNA before gel electrophoresis. (Lanes *1,4*) Contained 2 μl of control reticulocyte lysate; (lanes *1,2*) contained 1 μl of lysate from mock-transfected cells. The lower band in lane *3* (TFE3/Smad3 and Smad4) probably contains TFE3 and/or a Smad3–Smad4 complex that is bound to one of the two identical tandem repeats of the PE2.1 promoter fragment present in the probe.

supports this notion by showing that a complex of TFE3 and Smad3–Smad4 bind to the same PE2.1 element. TFE3 was generated by in vitro translation, and the complex of Smad4 and phosphorylated Smad 3 was produced in transfected cells. TFE3 (Fig. 10, lane 2) generates a single-shifted band, whereas the activated Smad3–Smad4 (Fig. 10, lane 4) generate two shifted bands. A distinct, slower-migrating band was detected in the sample containing both TFE3 and the activated Smad3–Smad4 complex (Fig. 10, lane 3). These results show that both TFE3 and the activated Smad3–Smad4 complex bind to the adjacent sequences of the same 36 nucleotide segment of the PE2.1 promoter.

Discussion

TFE3 activates TGF-β-induced transcription by binding to the E-box sequence in the PAI-1 *promoter*

To identify proteins that mediate TGF-β induction of the *PAI-1* promoter, we developed an expression cloning strategy utilizing an engineered TGF-β-responsive cell line and a retroviral cDNA library. Our strategy was based on the demonstration that the engineered TGF-β-responsive cell line, BAH–gpt, constructed by Hocevar and Howe (1996), grows in HAT medium only in the presence of TGF-β. In this sense, TGF-β was converted from a growth-inhibitory factor, its normal function, into a growth-promoting hormone. Our strategy also made use of the observation that retroviruses deliver recombinant DNA sequences into the genome of recipient cells at a very high efficiency (Kitamura et al. 1995), and that the ecotropic retrovirus receptor is essential and sufficient for infection of cells by murine retroviruses (Baker et al. 1992).

Using this approach, we cloned the transcription factor, TFE3, which slightly activates transcription of the *PAI-1* gene in the absence of TGF-β but strongly potentiates the ability of TGF-β to induce transcription (Figs. 3,4). TFE3 has not been implicated previously in TGF-β signaling. It was isolated previously by screening a phage expression library with a ^{32}P-labeled E box-containing sequence from the immunoglobulin heavy chain gene, and is ubiquitously expressed (Beckmann et al. 1990; Zhao et al. 1993).

Several lines of evidence suggest that TFE3 plays a critical role in activating TGF-β-dependent transcription of the *PAI-1* gene. First, cotransfection of TFE3 and a reporter gene containing ~800 bp of the natural *PAI-1* promoter enhances expression of the reporter gene fivefold in the presence of TGF-β (Fig. 3B). Second, serial truncation of the *PAI-1* promoter identified 36- to 56-bp segments that are responsive to both TGF-β addition and overexpression of TFE3 (Figs. 4, 5, 9C), and these elements contain an E-box sequence. Moreover, mutation of the E-box sequence in either the 56-bp PE2 promoter or the 36-bp PE2.1 promoter abolished TGF-β-induced transcription as well as its binding to TFE3 (Figs. 5 and 9C). Third, Smad3, a critical signal transducer in TGF-β signaling, synergizes with TFE3 in TGF-β-induced transcription (Fig. 6A) and mutation (Smad3A) of the TGF-β-inducible phosphorylation sites in Smad3 abolished its ability to activate transcription (Fig. 6B). Fourth, USF1, a basic helix–loop–helix transcription factor that also binds the E-box sequence (Beckman et al. 1992), activated transcription of a luciferase gene driven by the PE2.1 promoter in BAH-ER3 cells, but transcription of the reporter gene was no longer regulated by TGF-β (data not shown).

Phosphorylation of Smad3 triggers binding of a Smad3–Smad4 complex to a sequence in the PAI-1 *promoter adjacent to the TFE3-binding site*

Smad3 and Smad4 together, but neither alone, bind to the 36-bp PE2.1 promoter. A prerequisite for formation of this DNA-binding complex is that the cells express a constitutively active type I receptor TβRI–T204D (Fig. 7). The constitutively active TβRI phosphorylates Smad2 and Smad3, which are normally phosphorylated by the wild-type I receptor only on addition of TGF-β (Macias-Silva et al. 1996; Abdollah et al. 1997; Souchelnytskyi et al. 1997). Nevertheless, mutation of the TGF-β-inducible phosphorylation sites in Smad3 abolished the formation of a complex of Smad3 and Smad4 capable of binding to the PE2.1 sequence (Fig. 8).

Binding of the PE2.1 promoter by Smad3 and Smad4 was unaffected by mutation of the E box (data not shown), but was abrogated by mutation of the 7-bp sequence (5′-CCTAGAC-3′) located 3 bp upstream of the E box (Fig. 9). This suggests that at least part of the 5′-CCTAGAC-3′ sequence contains the Smad binding site. Dennler et al (1988) reported GST fusion proteins of both full-length Smad4 and the amino-terminal half of Smad3 directly bind to the CAGA sequence 5′ to the CCTAGAC sequence in the PE2.1 promoter. In contrast, our evidence suggests that phosphorylation of Smad3 not only triggers its association with Smad4, as reported previously (Nakao et al. 1997) but also is indispensable for binding to the PE2.1 element and subsequent activation of gene transcription (Figs. 6B and 9C). We have not yet precisely mapped the Smad binding site in this promoter segment. Phosphorylation may induce exposure of the DNA-binding domain in Smad3, or a multimer of Smad4 and phosphorylated Smad3 may have higher affinity for the PE2.1 sequence than does an unphosphorylated Smad3 monomer.

A number of recent reports show direct binding of Smad3 and Smad4 to specific DNA sequences, but these reports disagree on the consensus binding sequences (Yingling et al. 1997; Dennler et al. 1998; Vindevoghel et al. 1998; Zawel et al. 1998). As an example, *Drosophila* Mad binds to the consensus sequence GCCGnCGc (Kim et al. 1997); whereas human Smad 3 and Smad4 was reported to preferentially bind to GACACC (Yingling et al. 1997), GTCTAGAC (Zawel et al. 1998), or AG(C/A)CAGACA (Dennler et al. 1998); the latter sequence is also present in the PE2.1 element in the *PAI-1* promoter. Hence, Smad3 and Smad4 appear to bind to DNA with a

relative but not absolute specificity. Multiple tandem repeats of a Smad-binding sequence are required for TGF-β-induced transcription of a reporter gene, and even two tandem repeats of the GTCTAGAC sequence cannot support TGF-β-induced expression of a luciferase reporter gene (Zawel et al. 1998). These observations raise the possibility that a complex of multiple Smad proteins, together with other transcription factors such as TFE3, are required for maximal TGF-β-inducible transcription.

Synergism of TFE3 and Smad proteins in TGF-β-induced gene transcription by binding to adjacent sites in the PAI-1 *promoter*

TFE3 and a complex of Smad3 and Smad4 bind to adjacent sites in the 36-bp PE2.1 promoter (Figs. 9 and 10), and both binding sites are required for maximal TGF-β-induced gene transcription (Fig. 9C). This synergy requires TGF-β-induced phosphorylation of the carboxyl terminus of Smad3, as mutant Smad3A, lacking the TGF-β-induced phosphorylation sites, cannot synergize with TFE3 to activate TGF-β-dependent transcription from the PE2.1 promoter (Fig. 6B).

The model for TGF-β-induced transcription of the *PAI-1* gene in Figure 11 summarizes our results. A TGF-β-activated type I receptor phosphorylates Smad3, which then associates with Smad4. The complex of Smad3 and Smad4 then enters the nucleus and binds to a sequence upstream of the E box, which is already occupied by TFE3. Binding of both the Smad3–Smad4 complex and TFE3 within the 36-bp PE2.1 element is essential for maximal transcription of the *PAI-1* gene.

Our model differs somewhat from that proposed for activin-induced activation of the *Xenopus* transcription factor FAST-1. On addition of activin, a complex of Smad4 and phosphorylated Smad2 forms in the cytosol, translocates into the nucleus, binds FAST1, and then binds to a segment in the promoter of the developmentally regulated *Mix2* gene (X. Chen et al. 1996, 1997). In contrast, we have been unable to detect an interaction of TFE3 and the Smad3–Smad4 complex in the absence of DNA. It is possible that the Smad3–Smad4 complex does not directly bind to TFE3, even though they bind to adjacent sites of the PE2.1 promoter. Alternatively, TFE3 may form a complex with another, as yet unidentified, transcription factor, and that only this complex associates with the Smad3–Smad4 complex that is formed after TGF-β stimulation.

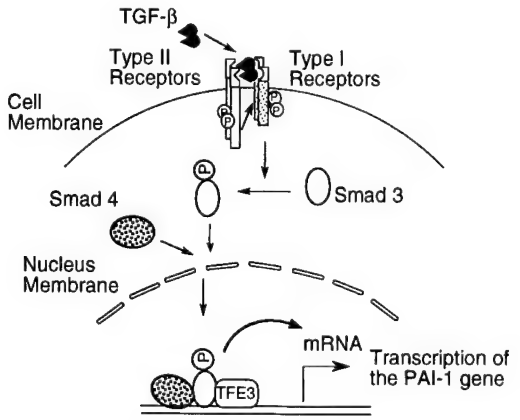

Figure 11. A model for cooperation of TFE3, Smad3, and Smad4 in TGF-β-induced activation of the *PAI-1* promoter. A complex of Smad4 and phosphorylated Smad3 bind to the DNA sequence 5′ to the E-box sequence, which is occupied by TFE3. Binding of the Smads potentiates the activity of TFE3, leading to TGF-β-induced transcription of *PAI-1* gene. Although indicated as monomers, we do not yet know the oligomeric state of the Smad or TFE3 proteins when they bind to the *PAI-1* promoter.

TGF-β activates a diverse range of genes in different cell types. Although Smad3–Smad4 complexes may bind specific DNA sequences, our work suggests that to induce expression of specific genes, these complexes need to cooperate with one or more transcription factors. We showed that TFE3 is one such factor essential for induction of the PE2.1 promoter. Because other genes activated by TGF-β do not have E boxes in the promoter sequenced to date, it is likely that other transcription factors interact with Smad3–Smad4 and Smad2–Smad4 complexes to induce transcription of other genes such as $p15^{INK4B}$ and $p21^{WAF1/CIP}$ (Datto et al. 1995; Li et al. 1995). The TFE3-binding sequence E box is essential for both basal and TGF-β-induced transcription of the PE2.1 promoter. We have no evidence to suggest that TFE3 itself is modified or activated following TGF-β addition; consistent with this notion, in vitro translated TFE3 binds to the E box in the PE2 promoter. We speculate that in unstimulated cells, TFE3 is bound to the two E boxes in the *PAI-1* promoter and supports a low level of transcription. Binding of an activated Smad3–Smad4 complex leads to a several-fold increase in gene expression.

Such synergistic cooperation of Smad3–Smad4 complexes with specific transcription factors offers the organism a distinct advantage—the same Smad3–Smad4 complex will activate different genes in different cells depending on the sequence of a promoter and the set of cooperative transcription factors that are expressed.

Materials and methods

Plasmid construction

Standard molecular biology techniques were used as described (Sambrook et al. 1989). Oligonucleotides were synthesized by GIBCO–BRL. To construct 3APP–Luc, the TATA box sequence, 5′-AGGGTATATAAT-3′, was inserted into the *Pst*I–*Bgl*II site of the pGL3–basic vector (Promega); a pair of oligonucleotides corresponding to −740/−644 of the *PAI-1* promoter was inserted upstream of the TATA box sequence; finally a pair of oligonucleotides containing three tandem repeats of the AP1-binding site [TGA(G/C)TCA] separated by an *Xba*I site was inserted upstream of the *PAI-1* promoter sequence. PAI–Luc was constructed by inserting the 0.8-kb *Hind*dIII fragment of the *PAI-1* promoter (Westerhausen et al. 1991) into the *Hin*dIII site of pGL3–basic. Various DNA fragments PCR-amplified from the *PAI-1* promoter were cloned into the *Kpn*I–*Pst*I sites of 3APP–Luc to generate PF1–Luc, PF2-Luc, and PF3-Luc in place of the 3APP promoter. These DNA fragments corresponded to nucleotides −794 to −532, −552 to −194, and −214 to +29 of the *PAI-1*

promoter (Keeton et al. 1991). Pairs of oligonucleotides including the E-box sequence in the *PAI-1* promoter were cloned into the *Kpn*I–*Pst* I sites of 3APP-Luc to generate PE1-Luc and PE2-Luc. These oligonucleotides corresponded to −740 to −644 and −583 to −528, respectively. To construct PE2.1–Luc, PE2.1S^m-Luc, and PE2.1E^m-Luc reporter genes, two tandem repeats of the wild-type or mutant oligonucleotides corresponding to −586 to −551 of the *PAI-1* promoter were inserted into *Kpn*I–*Pst*I sites of p3APP–Luc to replace its promoter (see Fig. 9A). All the constructs were sequenced to confirm the cloning junctions and more than one independent clone of each construct were tested in transfection for luciferase assays to confirm the results.

To generate a plasmid encoding Smad3, human Smad3 cDNA was amplified by PCR and cloned into the *Bam*HI–*Xba*I sites of a modified pcDNA3, resulting in a Smad3 fusion protein with two tandem repeats of a Flag epitope tag at its NH_2-terminus. A plasmid encoding Smad4 was generated by insertion of the human Smad4 cDNA into pMX–IRES–GFP (X. Liu et al. 1997). Human Smad4 cDNA was also cloned into the *Bam*HI–*Not*I sites of the vector pEXL, a derivative of pEGFP-N1 (Clontech) described previously (X. Liu et al. 1997) to generate pEXL–Smad4. Plasmids Flag-N–Smad3 and Flag-N–Smad3A were described previously (X. Liu et al. 1997). PCR-amplified TFE3 was cloned into the *Bam*HI–*Xho*I sites of pET28a (Novagen) to generate pET–TFE3 for in vitro transcription and translation. The kinase-defective mutant plasmid of the human TGF-β type I receptor, pCMV5–TβRI–KR, and the constitutively active TGF-β type I receptor, pCMV5–TβRI–T204D, were described previously (Wieser et al. 1995).

Tissue culture

BAH–gpt cells and HPRT-deficient HT1080 cells were kindly provided by P. Howe at the Cleveland Clinic Research Foundation (Hocevar and Howe 1996). HepG2 cells were purchased from ATCC. All of these cells were cultured in DMEM containing 10% fetal calf serum, 100 U/ml penicillin and 100 μg/ml streptomycin, in 5% CO_2 at 37°C unless otherwise stated. To treat cells with HAT or 6-TG, 1× HAT medium (GIBCO–BRL) or 30 μg/ml 6-TG (Sigma) was added to the normal medium. TGF-β1 was provided by R&D Systems, Inc., and a concentration of 200 pM was added to cell cultures as indicated.

Construction of a retroviral cDNA library and infection of cells by retroviruses

Poly(A)$^+$ RNA was isolated from HPRT-deficient HT1080 cells. cDNAs were synthesized from the poly(A)$^+$ RNA by use of the Superscript Plasmid System for cDNA Synthesis and Plasmid Cloning (GIBCO BRL) as described previously (Hua et al. 1996), and then cloned into the *Eco*RI–*Not*I sites of the retroviral vector pMX (Onishi et al. 1996). The resulting cDNA library was amplified in transformed bacteria and then introduced into a packaging cell line to obtain a high titer retroviral cDNA library. Briefly, Bosc23 cells, a cell line expressing the Gag, Pol, and Env proteins of Moloney Leukemia Virus (Pear et al. 1993), were seeded at a density of 2×10^6 cells per 60-mm dish in DMEM containing 10% fetal calf serum. On day 1, cells in each dish were transfected with 5 μg of the retroviral cDNA library in the presence of chloroquine (25 μM) to increase the virus titer. The transfected cells were switched to fresh medium 9 hr after transfection, and the supernatant containing the recombinant retroviruses was collected 48 hr after transfection.

Supernatant containing the retroviruses was incubated with the target cells for 6–9 hr in normal medium containing 4 μg/ml Polybrene (Sigma). To measure the titer of the library, pMX–LacZ-derived retroviruses were produced in parallel with the retroviral cDNA library and were used to infect BAH–ER3 cells or NIH 3T3 cells; infected cells were stained with X-gal for β-galactosidase expression. The titer of the retroviral cDNA library was deduced from that of the pMX–LacZ retroviruses produced in parallel.

To rescue recombinant retroviruses from infected BAH-ER3 cells, we first transfected Bosc23 cells with the plasmid pZAP (from D. Baltimore's laboratory, MIT, Cambridge, MA), which carries the entire cDNA sequence of the murine Moloney leukemia retrovirus genome. The supernatant containing the wild-type virus was collected 48 hr after transfection, and 1 ml of a 1:2 dilution of the supernatant was used to superinfect infected BAH-ER3 cells in six-well plates.

Transfection, luciferase assay, and preparation of cell lysates

Cells were transfected by the calcium phosphate precipitation method (Sambrook et al. 1989). For luciferase assays, cells were also transfected with 0.5 μg/well pSV-β or 0.2 μg/well of pCMV-β encoding the *lacZ* gene (Clontech) as an internal control to normalize the luciferase activity.

To transfect BAH-ER3 cells and Hep G2 cells, cells were seeded at a density of 50,000 cells/well in 12-well plates unless otherwise stated. On day 1, the cells were switched to fresh medium and then transfected by the calcium phosphate precipitation method. After overnight incubation, the cells were switched to normal medium and incubated for 6–8 hr. Afterward, serum-free medium with or without 200 pM TGF-β was added to the transfected cells; cells were harvested 20 hr after incubation with TGF-β for luciferase and β-galactosidase assays. The cells in each well of 12-well plates were lysed with 250 μl of 1× lysis buffer (Promega), and luciferase assays were carried out with 20 μl of cell lysates by use of the Luciferase Assay System (Promega) as detailed by the manufacturer. To normalize the luciferase activity, 20 μl of cell lysate was incubated with 100 μl of reaction buffer from the Luminescent B-galactosidase Detection Kit II (Clontech Laboratories, Inc.) as instructed by the manufacturer. Both the luciferase and β-galactosidase activities were measured by an AutoLumat LB953 luminometer (EG & G Berthold). All luciferase activities were normalized by the β-galactosidase activities and presented as an average from duplicate samples.

To obtain cell lysates for gel-shift assays, Bosc23 cells were transfected with desired plasmids and the total amount of DNA per 60 mm dish was adjusted to 7.5 μg by use of the plasmid pEXL–GFP. After overnight transfection, the cells were switched to normal medium, and harvested 24 hr later. Cells from each 60-mm dish were lysed in 150 μl of buffer containing the following components: 50 mM Tris at pH 8.0, 500 mM NaCl, 1% NP-40, 25 mM β-glycerophosphate, and 1× protease inhibitor cocktail Complete (Boehringer Mannheim). The lysed cells were rotated at 4°C at 60 rpm for 2 hr, and the supernatant was collected by centrifugation for use in gel-shift assays.

Gel-shift assay and immunoblotting

Gel-shift reactions were carried out in a total volume of 30 μl at room temperature. The components of the reaction buffer are as follows: 20 mM Tris at pH 8.0, 60 mM KCl, 0.7 mg/ml bovine serum albumin, 1 mM EDTA, 1.6 mM dithiothreitol, 1.6 mM $MgCl_2$, 0.3 % NP-40, 66 μg/ml poly(dI–dC)/poly(dI–dC) (Pharmacia), and 12% glycerol. Radiolabeled probes were made either by end labeling the annealed oligonucleotides with [γ-^{32}P]ATP or by PCR amplification in the presence of [α-^{32}P]dCTP. Briefly,

a pair of oligonucleotides corresponding to the PE2 fragment (−583 to −528) of the *PAI-1* promoter was end-labeled with [γ-^{32}P]ATP; alternatively, the two tandem repeats of the PE2.1 element (5′-CCTAGACAGACAAAACCTAGACAATCACGT-GGCTGG-3′), which comprise base pairs −586 to −551 of the human *PAI-1* promoter, were amplified by PCR from the reporter construct PE2.1–Luc in the presence of [α-^{32}P]dCTP.

The amplified probe was isolated on a native polyacrylamide gel as described previously (Wang et al. 1993), and 4×10^3 cpm was added to each reaction that had received the cell lysates 15 min earlier. One microliter of cell lysate (~30 μg of protein) was used in each reaction unless otherwise stated. For competition with wild-type or mutant oligonucleotides, a 50-fold molar excess of unlabeled oligonucleotides was added to the reaction buffer containing the cell lysate 15 min prior to addition of the ^{32}P-labeled probe. Twenty minutes after addition of the probe, the reaction was loaded onto a 4% polyacrylamide gel in 0.5× TBE buffer (Sambrook et al. 1989), and electrophoresis was carried out at 20 mA for 70 min. To supershift the DNA-binding activity with antibodies, 1 μl of the indicated antibody was added to each reaction; the reaction was loaded onto the gel after 15 min of incubation. All signals were detected on a Fujix BAS2000 PhosphorImager.

To detect the expression of Smad3 and Smad4 proteins in transfected Bosc23 cells, cell lysates prepared from the transfected Bosc23 cells were separated on 6%–18% gradient polyacrylamide gels and then transferred to Nitrocellulose blotting filters. The filters were blotted with 1 μg/ml anti-Flag (M2) antibody (Eastman Kodak) for detection of the Flag epitope-tagged Smad3, or with a 1:2000 dilution of an anti-Smad4 rabbit polyclonal antibody (Nakao et al. 1997). Bound primary antibodies were detected with horseradish peroxidase-labeled anti-mouse or anti-rabbit secondary antibodies, respectively, and developed with enhanced chemiluminescence reagents purchased from Pierce.

Acknowledgments

We thank Dr. P. Howe for kindly providing HPRT-deficient HT1080 cells and BAH–gpt cells. The cDNAs encoding human Smad2, Smad3, and Smad4 and TβRI were kind gifts from Dr. J. Massagué and Dr. R. Derynck. We thank Drs. C. Heldin and P. ten Dijke for providing the polyclonal antibody against human Smad4, and Dr. M. Tal for reagents for gel shift assays. TGF-β1 was a kind gift from R&D Systems, Inc. We also thank Drs. B. Schiemann and A. Sirotkin for reading the manuscript, and other members of the Lodish group for stimulating discussions. This work was supported by National Institutes of Health (NIH) grant CA63260 to H.F.L. X.H. was supported by a Damon Runyon–Walter Winchell Cancer Research Fund postdoctoral fellowship (DRG 1429) and X.L. was supported by a postdoctoral fellowship from the NIH.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked 'advertisement' in accordance with 18 USC section 1734 solely to indicate this fact.

References

Abdollah, S., M. Macias-Silva, T. Tzukazaki, H. Hayashi, L. Attisano, and J. Wrana. 1997. TβRI phosphorylation of Smad2 on Ser^{465} and Ser^{467} is required for Smad2–Smad4 complex formation and signaling. *J. Biol. Chem.* **272:** 27678–27685.

Attisano, L. and J. Wrana. 1998. Mads and Smads in TGFβ signaling. *Curr. Opin. Cell Biol.* **10:** 188–194.

Attisano, L., J.L. Wrana, F. Lopez-Casillas, and J. Massagué. 1994. TGF-β receptors and actions. *Biochimica et Biophysica Acta* **1222:** 71–80.

Baker, B.W., D. Boettiger, E. Spooncer, and J.D. Norton. 1992. Efficient retroviral-mediated gene transfer into human B lymphoblastoid cells expressing mouse ecotropic viral receptor. *Nucleic Acids Res.* **20:** 5234.

Beckmann, H., L.-K. Su, and T. Kadesch. 1990. TFE3: A helix-loop-helix protein that activates transcription through the immunoglobulin enhancer μE3 motif. *Genes & Dev.* **4:** 167–179.

Chen, X., M.J. Rubock, and M. Whitman. 1996. A transcriptional partner for MAD proteins in TGF-β signalling. *Nature* **383:** 691–696.

Chen, X., E. Weisberg, F. Fridmacher, M. Watanabe, G. Naco, and M. Whitman. 1997. Smad4 and FAST-1 in the assembly of activin-responsive factor. *Nature* **389:** 85–89.

Chen, Y., J.-J. Lebrun, and W. Vale. 1996. Regulation of transforming growth factor β- and activin-induced transcription by mammalian Mad Proteins. *Proc. Natl. Acad. Sci.* **93:** 12992–12997.

Datto, M.B., Y. Yu, and X.-F. Wang. 1995. Functional analysis of the transforming growth factor β responsive elements in the WAF1/Cip1/p21 promoter. *J. Biol. Chem.* **270:** 28623–28628.

Dennler, S., S. Itoh, D. Vivien, P. ten Dijke, S. Huet, and J. Gauthier. 1998. Direct binding of Smad3 and Smad4 to critical TGF-β-inducible elements in the promoter of human plasminogen activator inhibitor-type I gene. *EMBO J.* **17:** 3091–3100.

Franzen, P., P. ten Dijke, H. Ichijo, H. Yamashita, P. Schulz, C. Heldin, and K. Miyazono. 1993. Cloning of a TGF-β type I receptor that forms a heteromeric complex with the TGF-β type II receptor. *Cell* **75:** 681–692.

Hannon, G. and D. Beach. 1994. p15^{INK4B} is a potential effector of TGF-β-induced cell cycle arrest. *Nature* **371:** 257–261.

Heldin, C.-H., K. Miyazono, and P. ten Dijke. 1997. TGF-β signalling from cell membrane to nucleus through SMAD proteins. *Nature* **390:** 465–471.

Hocevar, B. and P. Howe. 1996. Isolation and characterization of mutant cell lines defective in transforming growth factor β signalling. *Proc. Natl. Acad. Sci.* **93:** 7655–7600.

Hua, X., A. Nohturfft, J.L. Goldstein, and M.S. Brown. 1996. Sterol resistance in CHO cells traced to point mutation in SREBP cleavage-activating protein. *Cell* **87:** 415–426.

Keeton, M.R., S.A. Curriden, A.-J.V. Zonneveld, and D.J. Loskutoff. 1991. Identification of regulatory sequences in the type 1 plasminogen activator inhibitor gene responsive to transforning growth factor-β. *J. Biol. Chem.* **266:** 23048–23052.

Kim, J., K. Johnson, H.J. Chen, S. Carroll, and A. Laughon. 1997. Drosophila Mad binds to DNA and directly mediates activation of vestigial by Decapentaplegic. *Nature* **388:** 304–308.

Kitamura, T., M. Onishi, S. Kinoshita, A. Shibuya, A. Miyajima, and G.P. Nolan. 1995. Efficient screening of retroviral cDNA expression libraries. *Proc. Natl. Acad. Sci.* **92:** 9146–9150.

Lagna, G., A. Hata, A. Hemmati-Brivanlou, and J. Massague. 1996. Partnership between DPC4 and SMAD proteins in TGF-β signalling pathways. *Nature* **383:** 832–836.

Li, J.-M., M.A. Nichols, S. Chandrasekharan, Y. Xiong, and X.-F. Wang. 1995. Transforming growth factor β activates the promoter of cyclin-dependent kinase inhibitor p15^{INK4B} through an Sp1 consensus site. *J. Biol. Chem.* **270:** 26750–26753.

Lin, H.Y. and H.F. Lodish. 1993. Receptors for the TGF-β su-

perfamily: Multiple polypeptides and serine/threonine kinases. *Trends Cell Biol.* **3:** 14–19.

Liu, F., C. Pouponnot, and J. Massague. 1997. Dual role of the Smad4/DPC4 tumor suppressor in TGF-β-inducible transcriptional complexes. *Genes & Dev.* **11:** 3157–3167.

Liu, X., Y. Sun, S.N. Constantinescu, E. Karam, R.A. Weinberg, and H.F. Lodish. 1997. Transforming growth factor β-induced phosphorylation of Smad3 is required for growth inhibition and transcriptional induction in epithelial cells. *Proc. Natl. Acad. Sci.* **94:** 10669–10674.

Luo, K. and H.F. Lodish. 1996. Signalling by chimeric erythropoietin-TGF-β receptors: Homodimerization of the cytoplasmic domain of the type I TGF-β receptor and heterodimerization with the type II receptor are both required for intracellular signal transduction. *EMBO J.* **15:** 4485–4496.

Macias-Silva, M., S. Abdollah, P.A. Hoodless, R. Pirone, L. Attisano, and J.L. Wrana. 1996. MADR2 is a substrate of the TGF-β receptor and its phosphorylation is required for nuclear accumulation and signaling. *Cell* **87:** 1215–1224.

Massague, J., A. Hata, and F. Liu. 1997. TGF-β signalling through the Smad pathway. *Trends Cell Biol.* **7:** 187–192.

Nakao, A., T. Imamura, S. Souchelnystkyi, M. Kawabata, A. Ishisaki, E. Oeda, K. Tamaki, J. Hanai, C.-H. Heldin, K. Miyazono, P. ten Dijke et al. 1997. TGF-β receptor-mediated signalling through Smad2, Smad3, and Smad4. *EMBO J.* **16:** 5353–5362.

Onishi, M., S. Kinoshita, Y. Morikawa, A. Shibuya, J. Phillips, L.L. Lanier, D.M. Gorman, G.P. Nolan, A. Miyajima, and T. Kitamura. 1996. Applications of retrovirus-mediated expression cloning. *Exp. Hematol.* **24:** 324–329.

Pear, W.S., G.P. Nolan, M.L. Scott, and D. Baltimore. 1993. Production of high-titer helper free retroviruses by transient transfection. *Proc. Natl. Acad. Sci.* **90:** 8392–8396.

Rasheed, S. 1995. Retroviruses and oncogenes. In *The retroviride* (ed. J.A. Levy), pp. 293–306. Plenum Press. New York, NY.

Riccio, A., P.V. Pedone, L.R. Lund, T. Olesen, H.S. Olsen, and P.A. Andreasen. 1992. Transforming growth factor β1-responsive element: Closely associated binding sites for USF and CCAAT-binding transcription factor-nuclear factor I in the type 1 plasminogen activator inhibitor gene. *Mol. Cell. Biol.* **12:** 1846–1855.

Sambrook, J., E. Fritsch, and T. Maniatis. 1989. *Molecular cloning.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Souchenlnytskyi, S., K. Tamaki, U. Engstrom, C. Wernstedt, P. ten Dijke, and C.-H. Heldin. 1997. Phosphorylation of Ser465 and Ser467 in the C terminus of Smads mediates interaction with Smad4 and is required for transforming growth factor-β signaling. *J. Biol.Chem.* **272:** 28107–28115.

Vindevoghel, L., A. Kon, R. Lechleider, J. Uitto, A. Roberts, and A. Mauviel. 1998. Smad-dependent transcriptional activation of human type VII collagen gene (COLL7A1) promoter by transforming growth factor-β. *J. Biol. Chem.* **273:** 13053–13057.

Wang, X., M.R. Briggs, X. Hua, C. Yokoyama, J.L. Goldstein, and M.S. Brown. 1993. Nuclear protein that binds sterol regulatory element of low density lipoprotein receptor promoter II. *J. Biol. Chem.* **286:** 14497–14504.

Weis-Garcia, F. and J. Massagué. 1996. Complementation between kinase-defective and activation-defective TGF-β receptors reveals a novel form of receptor cooperativity essential for signaling. *EMBO J.* **15:** 276–289.

Westerhausen Jr., D.R., W.E. Hopkins, and J.J. Billadello. 1991. Multiple transforming growth factor-β-inducible elements regulate expression of the plasmigogen activator inhibitor type-1 gene in Hep G2 cells. *J. Biol. Chem.* **266:** 1092–1100.

Wieser, R., J. Wrana, and J. Massague. 1995. GS domain mutations that constitutively activate TβR-I, the downstream signaling component in the TGF-β receptor complex. *EMBO J.* **14:** 2199–2208.

Wrana, J.L., L. Attisano, J. Carcamo, A. Zentella, J. Doody, M. Laiho, X.-F. Wang, and J. Massagué. 1992. TGFβ signals through a heteromeric protein kinase receptor complex. *Cell* **71:** 1003–1014.

Wrana, J.L., L. Attisano, R. Wieser, V. Francesc, and J. Massagué. 1994. Mechanism of activation of the TGF-β receptor. *Nature* **370:** 341–347.

Yingling, J., M. Datto, C. Wong, J. Frederick, N. Liberati, and X. Wang. 1997. Tumor suppressor Smad4 is a transforming growth factor β-inducible DNA binding protein. *Mol. Cell. Biol.* **17:** 7019–7028.

Zawel, L., J. Dai, P. Buckhaults, S. Zhou, K. Kinzler, B. Vogelstein, and S. Kern. 1998. Human Smad3 and Smad4 are sequence-specific transcription activators. *Mol. Cell* **1:** 611–617.

Zhang, Y., X.-H. Feng, R.-Y. Wu, and R. Derynck. 1996. Receptor-associated Mad homologues synergize as effectors of the TGF-β response. *Nature* **383:** 168–172.

Zhao, G.-Q., Q. Zhao, X. Zhou, M.-G. Mattei, and B. DE Combrugghe. 1993. TFEC, a basic helix-loop-helix protein, forms heterodimers with TFE3 and inhibits TFE3-dependent transcription activation. *Mol. Cell. Biol.* **13:** 4505–4512.

THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 274, No. 9, Issue of February 26, pp. 5716–5722, 1999
Printed in U.S.A.

Transforming Growth Factor-β Induces Formation of a Dithiothreitol-resistant Type I/Type II Receptor Complex in Live Cells*

(Received for publication, March 30, 1998, and in revised form, December 1, 1998)

Rebecca G. Wells,[a,b,c,d] Lilach Gilboa,[c,e,f] Yin Sun,[b,g] Xuedong Liu,[b,h] Yoav I. Henis,[e] and Harvey F. Lodish[b,i,j]

From [b]The Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, [e]Department of Neurobiochemistry, The George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv 69978, Israel, [a]Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02115, and [i]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Transforming growth factor-β (TGF-β) binds to and signals via two serine-threonine kinase receptors, the type I (TβRI) and type II (TβRII) receptors. We have used different and complementary techniques to study the physical nature and ligand dependence of the complex formed by TβRI and TβRII. Velocity centrifugation of endogenous receptors suggests that ligand-bound TβRI and TβRII form a heteromeric complex that is most likely a heterotetramer. Antibody-mediated immunofluorescence co-patching of epitope-tagged receptors provides the first evidence in live cells that TβRI·TβRII complex formation occurs at a low but measurable degree in the absence of ligand, increasing significantly after TGF-β binding. In addition, we demonstrate that pretreatment of cells with dithiothreitol, which inhibits the binding of TGF-β to TβRI, does not prevent formation of the TβRI·TβRII complex, but increases its sensitivity to detergent and prevents TGF-β-activated TβRI from phosphorylating Smad3 *in vitro*. This indicates that either a specific conformation of the TβRI·TβRII complex, disrupted by dithiothreitol, or direct binding of TGF-β to TβRI is required for signaling.

The transforming growth factor-β (TGF-β)[1] ligands are members of a large superfamily of cystine knot growth factors, which includes decapentaplegic (dpp) from *Drosophila melanogaster* as well as the Müllerian-inhibiting substance and the activins and bone morphogenetic proteins from mammals. The TGF-βs are important modulators of development, the extracellular matrix, and the immune response; they are potent growth inhibitors in many cell types, and their receptors and some downstream signaling elements are tumor suppressors (1–5).

TGF-β signals through the sequential activation of two serine-threonine kinase cell surface receptors (6–10), termed type I and type II (TβRI and TβRII). These two receptors physically associate to form a stable complex (8, 11). Several chimeric receptor systems have established that this complex is required for signaling (12–16). Whether it is preformed or ligand-induced is controversial. Isolation of a TβRI·TβRII complex from detergent lysates was possible only after pretreatment with TGF-β (9). Ligand-independent interactions between receptor cytoplasmic domains, however, have been detected in transfected COS cells and in the yeast two-hybrid system,[2] and work with TGF-β2 (which requires both receptors to bind) suggests that at least a small percentage of the cell surface receptor population is in preformed complexes (9, 17–19). Data from experiments on the effect of DTT also raise questions about the role of ligand in the complex. Treatment of cells with DTT prevents TGF-β binding to TβRI but not to TβRII (8, 20) and, *in vitro*, prevents formation of the TβRI·TβRII complex.[3]

The stoichiometry of the signaling complex is not known. TβRI and TβRII form ligand-independent homodimers in the endoplasmic reticulum and on the cell surface (21, 22). The simplest model is that two homodimers form a heterotetrameric signaling complex induced or activated by TGF-β. This view is supported by studies based on functionally complementary type I receptor mutants and on chimeric TGF-β/erythropoietin receptors demonstrating that type I dimers are required for signaling (13, 23). Yamashita *et al.* (24) used nonreducing/reducing two-dimensional SDS-PAGE to isolate ligand-bound TβRI and TβRII homo- and heterodimers. They speculate that these are derived from heterotetramers, although their data are also consistent with smaller complexes.

We report here studies on the physical nature of the TβRI·TβRII complex. We demonstrate that the complex is most likely a stable heterotetramer. Our studies show conclusively that some TβRI·TβRII complexes exist at the surface of live cells in the absence of ligand, and that TGF-β significantly enhances heterocomplex formation. DTT treatment, which prevents TGF-β binding to TβRI, does not inhibit complex forma-

* This work was supported by National Institutes of Health Grants CA63260 (to H. F. L.) and DK02290 (to R. G. W.) and by grants from the Israel Science Foundation administered by the Israel Academy of Arts and Sciences and from the Israel Cancer Research Fund (to Y. I. H.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[c] These authors contributed equally to this work.
[d] Current address: Dept. of Medicine, Yale School of Medicine, P. O. Box 208019, New Haven, CT 06520.
[f] Recipient of a fellowship from the Clore Scholars Programme.
[g] Supported by a postdoctoral fellowship from the Robert Steel Foundation for Pediatric Cancer Research.
[h] Supported by a National Institutes of Health postdoctoral fellowship.
[j] To whom correspondence should be addressed: The Whitehead Institute for Biomedical Research, 9 Cambridge Ctr., Cambridge, MA 02142. Tel.: 617-258-5216; Fax: 617-258-6768; E-mail: lodish@wi.mit.edu.

[1] The abbreviations used are: TGF-β, transforming growth factor-β; TβRI, TβRII, TβRIII, types I, II, and III TGF-β receptors; octyl-POE, *n*-octylpolyoxyethylene; DTT, dithiothreitol; PAGE, polyacrylamide gel electrophoresis; PBS, phosphate-buffered saline; MES, 4-morpholineethanesulfonic acid; GST, glutathione *S*-transferase; HA, hemagglutinin.

[2] R. Perlman and R. A. Weinberg, personal communication.
[3] C. Rodriguez, R. Lin, R. G. Wells, P. Scherer, and H. F. Lodish, manuscript in preparation.

tion but does result in the failure of TβRI from DTT-treated cells to phosphorylate Smad3 *in vitro*. We suggest that ligand binding to the two receptors may have different functions, and that complex formation itself is not sufficient for signal initiation.

EXPERIMENTAL PROCEDURES

Materials and Constructs—COS7 (CRL 1651), L6 (CRL 1458), and Mv1Lu (CCL 64) cells were grown as described previously (22, 25). 9E10 (α-Myc) mouse ascites was from Harvard Monoclonals and 12CA5 (α-HA) from BabCO. Fluorophore-labeled affinity-purified antibodies, Cy3-streptavidin, and biotinylated $F(ab')_2$ of GαM (goat anti-mouse $F(ab')_2$ were from Jackson Immunoresearch Laboratories. IgG fractions and monovalent F(ab') fragments were prepared as described (22, 26, 27). Untagged TβRI, TβRII, and N-terminally HA- and Myc-tagged receptors were as described previously (6, 10, 21, 22).

Binding and Cross-linking—Radioiodination of TGF-β1 (Celtrix Laboratories and R&D Systems) and binding and cross-linking of subconfluent cells were as described (25). Cells were preincubated (30 min, 37 °C) in KRH (50 mM HEPES, pH 7.5, 128 mM NaCl, 1.3 mM $CaCl_2$, 5 mM $MgSO_4$, 5 mM KCl) containing 0.5% fatty acid free BSA (KRH/BSA; Sigma). DTT-treated cells were incubated with 2 mM DTT (5 min, 37 °C) and then rinsed three times with warm KRH/BSA. Cells were then incubated (1–4 h, 4 °C) in fresh KRH/BSA with 100 pM ^{125}I-TGF-β1. Cross-linking was performed with 0.5 mg/ml disuccinimidyl suberate (Pierce) for 15 min, followed by quenching with 20 mM glycine. Cells to be used for gradients (Fig. 1) were then rinsed, incubated in 0.2 mM iodoacetamide in KRH (15 min, 4 °C), and lysed in 150 μl of MNT lysis buffer (20 mM MES, pH 6.0, 30 mM Tris, pH 7.4, 100 mM NaCl, with 2% *n*-octyl-polyoxyethylene (octyl-POE; Bachem Bioscience)). For co-immunoprecipitation (Fig. 3), cells were lysed in 1 ml of various lysis buffers (see Fig. 3 legend). After clearing the lysates, 100 μl were analyzed directly by SDS-PAGE. The remainder was split in half and immunoprecipitated (overnight, 4 °C) with either 10 μl/ml of antibody α-IIC, a polyclonal rabbit antiserum raised against the C-terminal 16 amino acids of the type II receptor (11), or 14 μl/ml of antibody VPN, raised against the juxtamembrane region of the human type I receptor (6). After an additional 30 min of incubation with 50 μl 1:1 protein A-Sepharose in PBS, bound beads were rinsed twice with the original lysis buffer then once with PBS. Protein was eluted into SDS-sample buffer and analyzed by SDS-PAGE.

Velocity Centrifugation on Sucrose Gradients—The velocity centrifugation technique has been described elsewhere (22). Briefly, cleared lysates (150 μl in MNT lysis buffer) were mixed with 50 μl size markers (29–669 kDa; Sigma) and layered over 7.5–30% sucrose gradients in MNT/1% octyl-POE. Centrifugation was for 8 h, 60,000 rpm, in an SW60 rotor (Beckman) at 4 °C. 250 μl fractions were removed sequentially from the top of each gradient. After removal of 25 μl for SDS-PAGE and Coomassie staining (to analyze migration of markers for each individual gradient), fractions were immunoprecipitated with α-IIC or VPN. Samples were analyzed by 7.5% SDS-PAGE. Autoradiographs were quantified with a LaCie Silverscanner II and MacBAS (Fuji) software.

Immunofluorescence Co-patching—The method used has been described elsewhere (21, 22). In the protocol used (detailed in Ref. 21), COS7 cells co-transfected with TβRI-Myc and TβRII-HA were preincubated (30 min, 37 °C) in serum-free Dulbecco's modified Eagle's medium, washed twice with cold Hanks' balanced salt solution with 20 mM HEPES, pH 7.4, containing 1% fatty acid-free BSA, and incubated successively with (*a*) normal goat IgG (200 μg/ml) to block nonspecific binding; (*b*) α-Myc F(ab') (50 μg/ml); (*c*) F(ab') of biotinylated GαM (5 μg/ml); (*d*) F(ab') of unlabeled GαM (200 μg/ml), used to block all free epitopes on the α-Myc F(ab'); (*e*) α-HA IgG (20 μg/ml); (*f*) fluorescein isothiocyanate-labeled GαM (20 μg/ml). After fixation (3.2% para-formaldehyde in PBS, pH 7.4, with 1.1% lysine and 0.24% $NaIO_4$) and quenching with 50 mM glycine in PBS (21), the cells were incubated with 0.5 μg/ml Cy3 streptavidin, and mounted with mowiol (Hoechst) containing 29 mM n-propyl gallate (Sigma). Note that membrane proteins retain some lateral mobility after paraformaldehyde fixation, and can therefore be patched by streptavidin because of its super-high binding affinity and multivalent nature. Fluorescence microscopy and digital image acquisition (CC/CE 200 CCD camera, Photometrics) were described (21). For each field, fluorescein and Cy3 images were taken separately using highly selective filter sets; the two images were superimposed, exported to Photoshop (Adobe) and printed (21).

GST-Smad3 Kinase Assay—The construction of cell line Mv1Lu-Flag-N-Smad3, Mv1Lu in which an N-terminally Flag-tagged Smad3 gene was stably expressed, was described elsewhere (28). GST-Smad3 (construct provided by Ying Zhang and Rik Derynck) (29) protein was prepared from bacterial lysate as described (30). Mv1Lu cells (Fig. 4*A*) were incubated in KRH/BSA for 30 min at 37 °C, followed by 5 min with or without 2 mM DTT. They were washed three times with KRH/BSA, and incubated with 100 pM unlabeled TGF-β1 (10 min, 37 °C). Cells were then washed and lysed in 1 ml of lysis buffer (150 mM NaCl, 1% Nonidet P-40, 50 mM Tris-HCl, pH 7.5, 50 mM NaF, 50 mM β-glycerophosphate, 1 mM Na_3VO_3, 1 mM DTT, 5 mM EDTA, pH 8.0, 1 mM phenylmethylsulfonyl fluoride, 1 mg/ml leupeptin, 10% glycerol). Supernatants were equalized for protein content and incubated with 2 μg of GST-Smad3 (3 h, 4 °C). The GST fusion protein was bound to 10 μl of glutathione beads that were washed three times with lysis buffer, then washed again with kinase buffer (50 mM NaCl, 20 mM Tris, pH 7.5, 12 mM $MgCl_2$, 5 mM DTT) and subjected to an *in vitro* kinase reaction with 0.2 mM unlabeled ATP and 50 μCi [γ-^{32}P]ATP (20 μl, 30 °C, 30 min). For Fig. 4*B*, lysates from Mv1Lu-Flag-N-Smad3 cells treated similarly were immunoprecipitated with the anti-TβRI antibody VPN (6), with or without 3 μg/ml competing peptide, before the incubation with GST-Smad3 and [γ-^{32}P]ATP under the above conditions. SDS-containing sample buffer was added, and the samples boiled to terminate the kinase reaction. Samples were analyzed by 8% SDS-PAGE.

RESULTS

TβRI and TβRII Form Stable, Heterotrimeric, or Heterotetrameric Complexes—We studied the size of the cell surface ligand-bound and cross-linked TβRI·TβRII complex using sucrose gradient velocity centrifugation, a technique we used previously to show that both receptors form homodimer-sized complexes in the endoplasmic reticulum (22). Cells were lysed with octyl-POE, a nonionic detergent with a density of 1 and a high CMC that enables analysis of the migration of detergent-solubilized proteins rather than of micelles (31). ^{125}I-TGF-β1-bound and cross-linked TβRII from COS7 cells transfected with TβRII alone migrated with a peak centered at fraction 7, the position of a 150-kDa marker protein (Fig. 1*A*, *top* and *bottom panels*). Although absolute size determinations can be inaccurate for detergent-solubilized membrane proteins, this is consistent with a homodimer bound to one or more molecules of TGF-β1, as expected from previous gradient analysis and immunofluorescence (21, 22). TβRII from COS7 cells cotransfected with both TβRI and TβRII (Fig. 1*A*, *middle* and *bottom panels*) and immunoprecipitated with an antibody against TβRI, thus in a TβRI·TβRII complex, migrated more quickly with a peak centered around fraction 8, indicative of a heterotrimeric or tetrameric complex.

L6 rat myoblasts, TGF-β-responsive cells which lack the type III TGF-β receptor, show similar receptor complexes. TGF-β1-bound type I and II receptors immunoprecipitated with an antibody against TβRII both migrated with a peak centered at fraction 10, correlating with markers between 230 and 265 kDa (Fig. 1*B*, *top* and *bottom panels*), and consistent with a heterotetrameric or trimeric complex. Nearly identical results were obtained with TGF-β2 and with other cell lines, including Mv1Lu, HYB2, THP1, and GH3 (data not shown). Immunoprecipitation with an antibody against TβRI (data not shown) gave similar results, indicating the absence of a significant population of ligand-bound TβRII, which is not in a complex with TβRI. Pretreatment of cells with DTT (Fig. 1*B*, *middle* and *bottom panels*) prevented ligand binding and cross-linking to TβRI and slowed the migration of TβRII to a peak centered around 155 kDa. The DTT-treated complex is likely a dimer (for comparison, see singly transfected COS cells in Fig. 1*A*), indicating disruption of the TβRI·TβRII complex under these conditions. DTT does not appear to affect TβRII homodimers, a finding consistent with previous results (21). The DTT effect appears to be on the cross-linking of ligand to TβRI rather than on cross-linking between TβRI and TβRII, because immunoprecipitation of TβRI from DTT-treated cells following ligand binding and cross-linking showed the absence of ligand-labeled

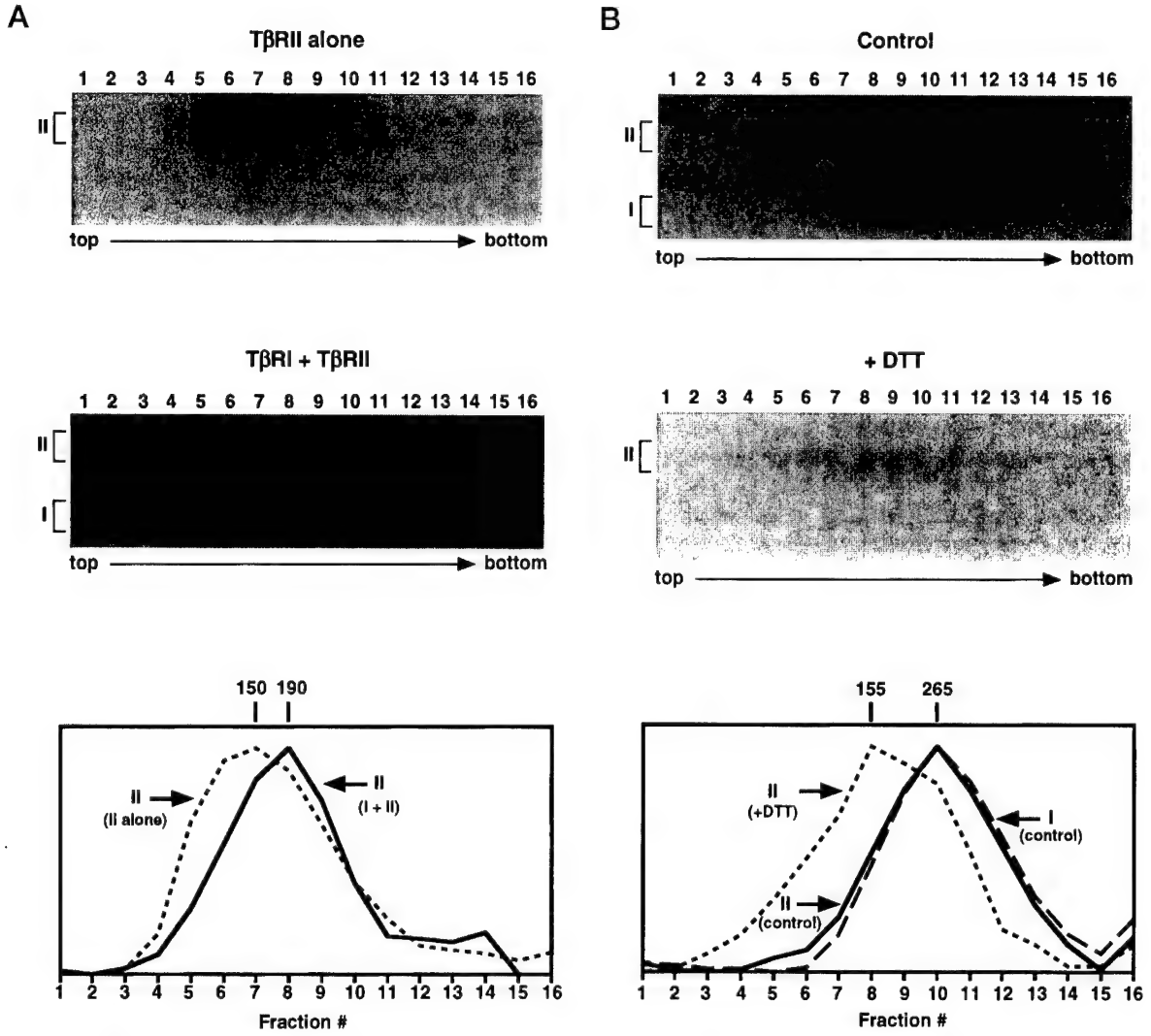

FIG. 1. **Type I and II receptors form stable heterotetrameric or trimeric complexes in COS7 and L6 cells.** Cells were cross-linked to ^{125}I-TGF-β1, lysed, and subjected to velocity centrifugation over sucrose gradients. *A*, COS7 cells were transfected with TβRII alone (*top panel*) or together with TβRI (*middle panel*). Fractions from the sucrose gradients were immunoprecipitated with anti-TβRII (*top*) or anti-TβRI (*middle*) and analyzed by SDS-PAGE. Migration of the type II receptor on each gradient was quantified and normalized to the peak value and is shown in the *bottom panel*. Size estimates (in kDa) are from size markers included in each gradient tube. Receptors from doubly transfected cells (*middle*) were immunoprecipitated with antibody against TβRI, ensuring that all TβRII measured is part of the complex and not from singly transfected cells. Quantification of the immunoprecipitated ligand-labeled type I receptor results in a nearly identical curve (data not shown). The broad nature of the peak for cells transfected with TβRII alone may reflect some association of TβRII with the small number of endogenous TβRI found in COS7 cells. *B*, L6 cells were treated as in *A*, except that cells in the *middle panel* were treated with 2 mM DTT before ligand binding. Fractions were immunoprecipitated with anti-TβRII. Migration of TβRII (*middle panel*) or both receptors (*top*) on the gradients was quantified (*bottom panel*). Note that fraction numbers are comparable only within a given experiment, and that migration of markers for part *A* and part *B* was slightly different. Control experiments demonstrate that only minimal dissociation of receptor-bound but noncross-linked ligand occurred in detergent lysates over the 8 h required for centrifugation (data not shown).

TβRI (Fig. 3). Analogously, ligand binding and cross-linking after DTT treatment either to endogenous TGF-β receptors (20) or to COS cells transiently expressing TβRI and TβRII (not shown) failed to reveal TβRI labeling.

TβRI and TβRII Form a Ligand-dependent Complex in Live Cells—To study the TβRI·TβRII complex in live cells, we used TβRI and TβRII carrying HA or Myc epitope tags at their extracellular termini for immunofluorescence co-patching, a technique that we developed and have described previously (21, 22). Briefly, a tagged receptor at the surface of live, unfixed cells (in the cold to avoid internalization) is forced into patches by a double layer of bivalent IgGs where the secondary antibody is coupled to one fluorophore (*e.g.* fluorescein, which emits green fluorescence). A second receptor, containing a different tag, is labeled by antibodies coupled to a second fluorophore (*e.g.* Cy3, red fluorescence). The cells are examined by fluorescence microscopy to determine whether the two receptors are swept into mutual (yellow) or separate (red or green) micropatches. We have employed this method successfully to demonstrate that all three TGF-β receptors form ligand-independent homodimers (21, 22).

Fig. 2 shows the results of co-patching experiments performed on COS7 cells co-transfected with TβRI-Myc and TβRII-HA. The labeling specificity is high, as shown in a control experiment on cells transfected with TβRI-Myc alone (Fig. 2*A*, *inset*); these cells show only Cy3 labeling. In the absence of ligand (Fig. 2*A*), 15–20% of the patches were mutual (yellow), although the majority were separate (either green or red). In the presence of 250 pM TGF-β1, the percentage of mutual patches markedly increased to 40–50% (Fig. 2*B*), demonstrating that TβRI and TβRII at the surface of live cells have an inherently low probability of forming heterocomplexes that is significantly enhanced by ligand binding. The fraction of a given receptor type in heterocomplexes is proportional to the

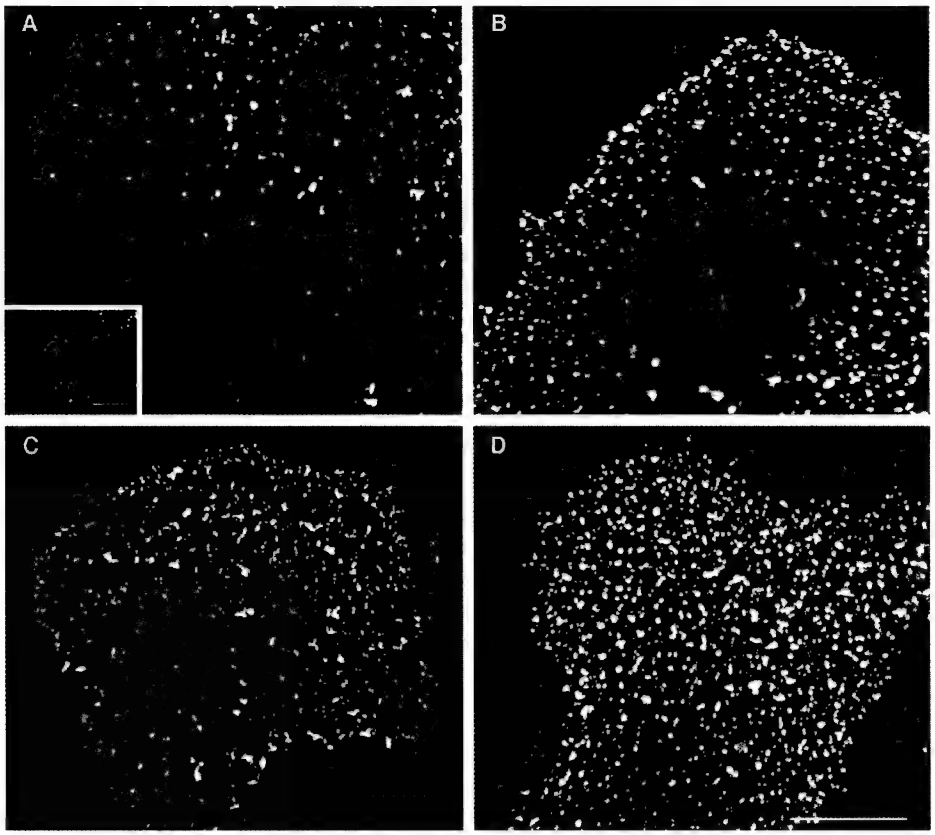

FIG. 2. **Type I/II hetero-oligomers are TGF-β-dependent and DTT-independent in live cells.** COS7 cells were co-transfected with TβRI-Myc and TβRII-HA (*A–D*), or with TβRI-Myc alone (*panel A, inset*). 48 h after transfection, cells in (*C*) and (*D*) were pretreated with 2 mM DTT (15 min, 37 °C). The live cells were incubated in the cold with (*B* and *D*) or without (*A* and *C*) 250 pM TGF-β1 for 2 h, followed by successive incubations with a series of antibodies to mediate patching and fluorescent labeling (see "Experimental Procedures"). The labeling protocol results in TβRI-Myc labeled by Cy3 (red), TβRII-HA labeled by fluorescein (green), and mutual patches containing both receptors labeled yellow upon superposition of the two fluorescent images. Cells transfected with TβRI-Myc alone (*panel A, inset*) are labeled exclusively with Cy3, demonstrating the labeling specificity. The numbers of red, green, and yellow patches were counted on the computer screen on 20 × 20 μm^2 flat cell regions (avoiding the nucleus, which contains more nonspecific staining and is out of the focal plane) for several independent experiments. Bars, 20 μm.

number of yellow patches divided by the sum of yellow and red (for the red-labeled receptor type) or yellow and green (for green-labeled receptors). These fractions are similar for TβRI and TβRII (23–29% and 57–63% in the absence and presence of ligand, respectively), in accord with a 1:1 stoichiometric ratio in the heterocomplex.

Ligand binding and cross-linking to TβRI co-expressed with TβRII is abrogated by DTT pretreatment of cells (Ref. 20 and Fig. 1*B*, *middle panel*). DTT treatment, however, did not dissociate the ligand-independent TβRI·TβRII complexes (around 20% co-patching), and did not affect the enhancement of heterocomplex formation by TGF-β1 (around 50% co-patching) (Fig. 2, *C* and *D*). The fraction of each receptor type in yellow patches also remained similar.

DTT Increases the Detergent Sensitivity of the Complex—In contrast to the persistence of TβRI·TβRII heterocomplexes in DTT-treated live cells (Fig. 2, *C* and *D*), velocity sedimentation experiments showed that receptor heterocomplexes were disrupted by DTT (Fig. 1*B*). We suspected that this disparity resulted from destabilizing effects of the detergent used to solubilize the receptors for velocity sedimentation. We therefore examined the effect of various detergents on TβRI·TβRII co-immunoprecipitation from DTT-pretreated Mv1Lu cells (Fig. 3). Immunoprecipitation with an antibody against TβRI (Fig. 3, *right panel*) was used to assess the integrity of the receptor complex in DTT-treated cells, because TβRII residing in heterocomplexes would still be labeled by ligand and would co-precipitate with TβRI even if the latter was unlabeled. Under the most stringent lysis conditions used (buffers 1–3), there was no co-immunoprecipitation, indicating the absence of intact heteromeric complexes after detergent solubilization. These results hold in different cell lines and with antibodies raised against different type I receptor epitopes (not shown). The same was also true with buffer 4, which was used for the velocity centrifugation experiment in Fig. 1*B*. With a fifth lysis buffer, however, we detected a small amount of intact TβRI·TβRII complex in the presence of DTT (Fig. 3, *right panel*, buffer 5), suggesting that buffer and detergent conditions determine the integrity of the complex in DTT-treated cells. Together with the demonstration of TβRI·TβRII heterocomplex formation in live cells pretreated with DTT (Fig. 2), these results suggest that DTT treatment alters the conformation of the heterocomplex formed; the altered complex is less stable (as reflected in its increased detergent sensitivity) and most likely has a different conformation, resulting in the failure of ligand binding and cross-linking to TβRI.

DTT Pretreatment Prevents TGF-β-induced Activation of the Type I Receptor—In epithelial cells, an essential step in TGF-β-mediated growth inhibition and PAI-1 promoter activation is activation of the ability of TβRI to rapidly phosphorylate Smad3 at its C-terminal SSVS motif (28). Because DTT treatment did not block the ligand-mediated association of TβRI with TβRII, we examined the signaling capability of these complexes as reflected by their ability to phosphorylate Smad3 *in vitro*. Mv1Lu cells (untreated or pretreated with DTT) were exposed to TGF-β1 to allow heterocomplex formation and activation of TβRI. In one study, lysates from these cells were incubated with recombinant GST-Smad3, and the isolated complex was subjected to an *in vitro* kinase reaction. In untreated cells (Fig. 4*A*, *lanes 1* and *2*), TGF-β mediated a greater than

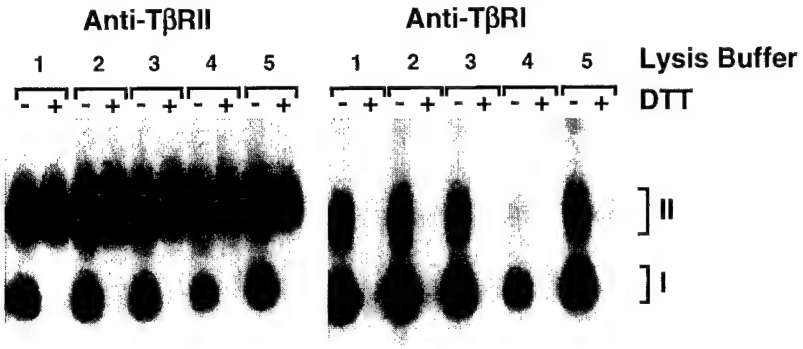

FIG. 3. **Detergent susceptibility of TβRI·TβRII receptor complexes in DTT-treated cells.** MV1Lu cells pretreated (+) or not (−) with 2 mM DTT were subjected to binding and cross-linking with 100 pM ^{125}I-TGF-β1. Cells were then lysed in various lysis buffers: *lane 1*, 0.5% Triton X-100, 0.5% deoxycholic acid, 10 mM EDTA in PBS; *lane 2*, 0.75% Triton X-100, 0.5% deoxycholic acid, 10 mM EDTA in PBS; *lane 3*, 1% Triton X-100, 0.5% deoxycholic acid, 10 mM EDTA in PBS; *lane 4*, 2% octyl-POE in MNT; *lane 5*, 0.5% Triton X-100, 150 mM NaCl, 1 mM EDTA, 20 mM Tris, pH 7.4. One-tenth of each lysate was analyzed by SDS-PAGE without further treatment (not shown); one-half of the remainder was immunoprecipitated with anti-TβRII (α-IIC), and the other half with anti-TβRI (VPN).

20-fold increase in Smad3 phosphorylation; two-dimensional tryptic mapping indicated that this *in vitro* phophorylation occurred at the same site as *in vivo* (data not shown). DTT-pretreated cells, however, showed no ligand-induced phosphorylation (Fig. 4*A*, *lanes 3* and *4*). Similar results (Fig. 4*B*) were obtained by first immunoprecipitating lysates with antibodies against TβRI, then incubating the complexes with a GST-Smad3 fusion protein in an *in vitro* kinase reaction. Although there is background phosphorylation of the GST-Smad3 construct in this assay, peptide competition during the immunoprecipitation (Fig. 4*B*, *lanes 5–8*) eliminates all of the TGF-β-inducible phosphorylation, indicating that TβRI is responsible. Thus, the TβRI·TβRII complex formed after DTT treatment is inactive and cannot mediate the earliest step in TGF-β downstream signaling phosphorylation of Smad3.

DISCUSSION

Our major findings are: 1) TβRI and TβRII form a stable, heteromeric complex, most likely a heterotetramer; 2) in live cells, the two receptors have an intrinsic affinity for each other that is markedly increased by TGF-β exposure; and 3) DTT does not prevent the formation of this complex but increases its detergent sensitivity and blocks TGF-β-induced activation of the type I receptor, as measured by its ability to phosphorylate Smad3.

We previously used velocity centrifugation to demonstrate that TβRI and TβRII each form homodimer-sized complexes in the endoplasmic reticulum (22). We use similar technology here to show that ligand-bound and cross-linked receptors in transfected COS7 cells and in L6 cells expressing native receptors form detergent-stable complexes whose size is consistent with heterotrimers or heterotetramers. In L6 cells (Fig. 1*B*), the migration of the TβRI·TβRII complex is consistent with a heterotetramer or heterotrimer bound to one TGF-β molecule. In COS7 cells, the complex migrates significantly faster than a homodimeric complex (Fig. 1*A*, *top panel*) (22); the wide nature of the peak, and the shoulder at higher fractions, may be because of dissociation of larger complexes during centrifugation. Furthermore, co-patching results obtained in live cells (Fig. 2) are most consistent with a heterotetramer, the fractions of TβRI and TβRII in mutual patches are similar, as expected for a stoichiometric ratio of 1:1. It should be stressed, however, that a definite determination of the stoichiometric ratio depends on an accurate measurement of the surface levels of both receptors on the cells scored, which is not feasible by the current methods.

The tetrameric nature of the complex is also supported by data indicating that a dimer of TβRII and more than one TβRI reside in the heterocomplex. Evidence that there are two TβRII in the complex comes from the demonstration by velocity centrifugation that treatment of cells with DTT results in a homodimer-sized complex of TβRII. This suggests that there were two type II receptors in the original, heteromeric complex; otherwise, one must assume that the normally dimeric TβRII (21, 22) is monomeric in the heterocomplex and reassociates to form dimers after DTT treatment and detergent solubilization. Type I receptors, which are homodimers when expressed alone and remain dimeric after DTT treatment, (22) are likely to be multimeric in the active complex, as indicated by experiments demonstrating the existence of functionally complementary TβRI mutants (23). Taken together, the sedimentation velocity, co-patching, and functional complementation studies imply that the signaling TGF-β receptor complex contains two TβRI and two TβRII polypeptides.

We provide the first evidence in live cells that the formation of the TβRI·TβRII complex is TGF-β-dependent (Fig. 2, *A* and *B*). Previous studies have demonstrated the TGF-β dependence of the TβRI·TβRII heterocomplex but used receptors in detergent lysates, leaving open the possibility that TGF-β stabilized the complex in detergent but did not induce its formation (8, 9, 11). Co-patching studies examining heterocomplex formation in the intact plasma membrane show a marked increase in mutual aggregates in the presence of TGF-β1. There is, however, an intrinsic affinity between the two receptors, as demonstrated by the 15–20% that reside in mutual aggregates in the absence of TGF-β (Fig. 2*A*). Although these experiments used cells overexpressing the transfected receptors, as required for visualization by immunofluorescence, there are several indications that the heterocomplexes observed are not the result of high expression levels. The co-patching experiments examine single cells under the microscope, and the use of double labeling by IgG or biotin/streptavidin enhancement enabled us to analyze cells expressing as few as 15,000 surface receptors (evaluated by photomultiplier-based measurement of the fluorescence intensity on cells expressing known receptor levels, as described by Henis *et al.* (21)). These cells, with receptor levels higher but of the same order of magnitude as untransfected cells, yielded the same co-patching results as cells expressing receptor levels 10-fold higher. In addition, in all cases where the level and percentage of higher complexes formed were high enough to allow detection by ultracentrifugation in untransfected cells expressing the receptors (including ligand-labeled TβRI·TβRII heterocomplexes in the current work, and TβRI or

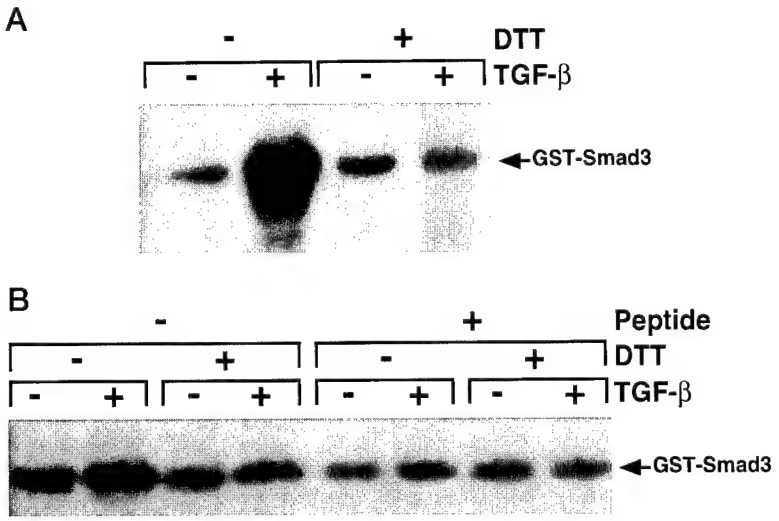

FIG. 4. **DTT pretreatment blocks TGF-β-mediated Smad3 phosphorylation.** *A*, Mv1Lu cells were treated with (*lanes 3* and *4*) or without (*lanes 1* and *2*) 2 mM DTT (5 min, 37 °C). After extensive washing, they were incubated with (*lanes 2* and *4*) or without (*lanes 1* and *3*) 100 pM TGF-β1 (10 min, 37 °C). Cells were rapidly chilled and lysed. Cleared lysates were incubated with a GST-Smad3 fusion protein for 3 h. After retrieval of the fusion protein by glutathione beads, it was subjected to an *in vitro* kinase reaction then analyzed by SDS-PAGE. Control cells were allowed to bind ^{125}I-TGF-β1 under the same conditions and were then cross-linked, lysed, and analyzed by SDS-PAGE; there was no binding of ligand to TβRI in DTT-pretreated cells, confirming that the 10-min incubation period at 37 °C was not sufficient for arrival of significant amounts of non-DTT-exposed TβRI at the cell surface (data not shown). *B*, Mv1Lu-Flag-N-Smad3 cells were treated as above, except lysates were immunoprecipitated with an anti-type I antibody (6) with (*lanes 5–8*) or without (*lanes 1–4*) 3 μg of competing peptide before incubation with the GST-Smad3 fusion protein.

TβRII homodimers in earlier studies) (22), a good agreement was obtained between the results on these cells and on transiently expressing COS cells. Furthermore, the results obtained in co-patching experiments depended on the receptor types examined, with low co-patching for type II/type III TGF-β receptor heterodimers *versus* high co-patching levels for TβRI and TβRII homodimers (21, 22). This emphasizes the specificity of the interactions measured. We note, however, that COS cells overexpressing TβRI and TβRII exhibit some ligand-independent receptor phosphorylation (18), raising the possibility that the number of TβRI·TβRII complexes is larger in this system than in cells expressing the receptors endogenously.

The ligand-independent association of a fraction of TβRI and TβRII is consistent with our previous finding that TGF-β2, unlike TGF-β1, binds to a preformed complex of TβRI and TβRII (19). Such preformed complexes do not appear to mediate TGF-β-independent signal transduction,[4] raising the question whether the role of ligand is to increase the number of complexes or to effect a necessary and stabilizing conformational change. A recent model of the TβRI ectodomain (based on certain cysteine motifs shared with protectin (CD59)) proposes that its surface has two distinct binding sites, one each for ligand and TβRII (32). This model is in agreement with the finding shown in Fig. 2 that there is an intrinsic affinity between TβRI and TβRII, which is increased by TGF-β.

To further investigate the relationship between TGF-β-mediated heterocomplex formation and signaling, we studied the effects of DTT treatment on complex stability and its ability to phosphorylate Smad3 *in vitro*. Pretreatment with DTT prevents TGF-β binding and cross-linking to the type I receptor (8, 20). We demonstrate that it also increases the detergent susceptibility of the complex (Figs. 1*B* and 3), and blocks TGF-β-induced activation of the ability of TβRI to phosphorylate Smad3 *in vitro* (Fig. 4) but does not change the percentage of TβRI and TβRII in mutual complexes with or without TGF-β1 (Fig. 2). The observation that multiple cross-linkers, including difluorodinitrobenzene (with a spacer length of 3Å) can cross-link TGF-β to TβRI suggests that cross-linking results reflect direct binding of ligand to TβRI rather than their fortuitous proximity within a complex.[4] Our DTT data therefore suggest that either a specific TβRI·TβRII complex conformation, destroyed by DTT, or direct ligand binding to TβRI are required for signal transduction. It is unclear whether ligand binding to TβRI stabilizes the complex with TβRII or serves a different function altogether.

Our data agree only in part with a previously published report that DTT pretreatment does not affect type I receptor phosphorylation and does not measurably alter the amount of ligand-bound and cross-linked TβRI·TβRII complex in detergent lysates (33). Although we found that TGF-β1 can mediate heterocomplex formation after DTT treatment, especially in intact cells (Fig. 2), we have been able to demonstrate at most a small percentage of TβRI·TβRII complex using the same detergent lysis conditions as these investigators (see Fig. 3, lysis buffer 5), in accord with a second report (6). In some of the experiments by Vivien and Wrana (33), the antibody used was against a C-terminal epitope of TβRI that has significant sequence identity to other type I receptors (34); however, antibody cross-reactivity could not explain all the differences between their results and ours.

These results raise interesting questions about the stoichiometry of TGF-β receptors in the signaling complex. Although our data are most consistent with a heterotetrameric $(T\beta RI)_2(T\beta RII)_2$ receptor complex, and we have obtained similar results with several cell lines, there are suggestions in the literature that the stoichiometry of the complex may vary. For example, the ratio of TβRI to TβRII in microvascular endothelial cells was significantly higher in three-dimensional *versus* two-dimensional cultures, corresponding to increasing resistance to the anti-proliferative but not the matrix-inducing effects of TGF-β (35). In other cell types, including bone, the ratio of type I to type II receptors may also be important, particularly in differentiating growth inhibition from other effects of TGF-β (36–38). The migration over sucrose gradients of the TβRI·TβRII complex from nonepithelial cell systems remains to be determined.

The number of ligand molecules in the complex is not known. The most likely possibilities are one TGF-β homodimer, each

[4] R. G. Wells and H. F. Lodish, unpublished results.

subunit bound to a type I and a type II receptor, or two ligand homodimers, each subunit bound to one of the four receptors in a presumed heterotetramer. The sucrose gradient data presented here show a surprisingly small increase in complex size when TβRI is added to TGF-β-bound TβRII, suggesting that the addition of TβRI to TβRII is not accompanied by the recruitment of additional ligand molecules.

Acknowledgments—TGF-β1 was a kind gift of R&D Biosystems. We are grateful to Ying Zhang and Rik Derynck (UCSF) for the GST-Smad3 construct and to Ralph Lin for comments on the manuscript.

REFERENCES

1. Roberts, A. B., and Sporn, M. B. (1990) in *Peptide Growth Factors and Their Receptors I* (Sporn, M. B., and Roberts, A. B., eds) Vol. 95/I, pp. 419–472, Springer-Verlag, Berlin Heidelberg
2. Kingsley, D. M. (1994) *Genes Dev.* **8,** 133–146
3. Attisano, L., Wrana, J. L., Lopez-Casillas, F., and Massague, J. (1994) *Biochim. Biophys. Acta.* **1222,** 71–80
4. Heldin, C.-H., Miyazono, K., and ten Dijke, P. (1997) *Nature* **390,** 465–471
5. Massague, J., and Weis-Garcia, F. (1996) *Cancer Surv.* **27,** 41–64
6. Franzen, P., ten Dijke, P., Ichijo, H., Yamashita, H., Schulz, P., Heldin, C. H., and Miyazono, K. (1993) *Cell* **75,** 681–692
7. Laiho, M., Weis, F. M. B., Boyd, F. T., Ignotz, R. A., and Massague, J. (1991) *J. Biol. Chem.* **266,** 9108–9112
8. Wrana, J. L., Attisano, L., Carcamo, J., Zentella, A., Doody, J., Laiho, M., Wang, X. F., and Massague, J. (1992) *Cell* **71,** 1003–1014
9. Wrana, J. L., Attisano, L., Wieser, R., Ventura, F., and Massague, J. (1994) *Nature* **370,** 341–347
10. Lin, H. Y., Wang, X. F., Ng-Eaton, E., Weinberg, R. A., and Lodish, H. F. (1992) *Cell* **68,** 775–785
11. Moustakas, A., Lin, H. Y., Henis, Y. I., Plamondon, J., O'Connor-McCourt, M. D., and Lodish, H. F. (1993) *J. Biol. Chem.* **268,** 22215–22218
12. Vivien, D., Attisano, L., Wrana, J. L., and Massague, J. (1995) *J. Biol. Chem.* **270,** 7134–7141
13. Luo, K., and Lodish, H. F. (1996) *EMBO J.* **15,** 4485–4496
14. Okadome, T., Yamashita, H., Franzen, P., Moren, A., Heldin, C. H., and Miyazono, K. (1994) *J. Biol. Chem.* **269,** 30753–30756
15. Anders, R. A., and Leof, E. B. (1996) *J. Biol. Chem.* **271,** 21758–21766
16. Muramatsu, M., Yan, J., Eto, K., Tomoda, T., Yamada, R., and Arai, K. (1997) *Mol. Biol. Cell* **8,** 469–480
17. Ventura, F., Doody, J., Liu, F., Wrana, J. L., and Massague, J. (1994) *EMBO J.* **13,** 5581–5589
18. Chen, F., and Weinberg, R. A. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 1565–1569
19. Rodriguez, C., Chen, F., Weinberg, R. A., and Lodish, H. F. (1995) *J. Biol. Chem.* **270,** 15919–15922
20. Cheifetz, S., and Massague, J. (1991) *J. Biol. Chem.* **266,** 20767–20772
21. Henis, Y. I., Moustakas, A., Lin, H. Y., and Lodish, H. F. (1994) *J. Cell Biol.* **126,** 139–154
22. Gilboa, L., Wells, R. G., Lodish, H. F., and Henis, Y. I. (1998) *J. Cell Biol.* **140,** 767–777
23. Weis-Garcia, F., and Massague, J. (1996) *EMBO J.* **15,** 276–289
24. Yamashita, H., ten Dijke, P., Franzen, P., Miyazono, K., and Heldin, C. H. (1994) *J. Biol. Chem.* **269,** 20172–20178
25. Wells, R. G., Yankelev, H., Lin, H. Y., and Lodish, H. F. (1997) *J. Biol. Chem.* **272,** 11444–11451
26. Henis, Y. I., Gutman, O., and Loyter, A. (1985) *Exp. Cell Res.* **160,** 514–526
27. Kurkela, R., Vuolas, L., and Vihko, P. (1988) *J. Immunol. Methods* **110,** 229–236
28. Liu, X., Sun, Y., Constantinescu, S. N., Karam, E., Weinberg, R. A., and Lodish, H. F. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 10669–10674
29. Zhang, Y., Feng, X.-H., Wu, R.-Y., and Derynck, R. (1996) *Nature* **383,** 168–172
30. Guan, K. L., and Dixon, J. E. (1991) *Anal. Biochem.* **192,** 262–267
31. Garavito, R. M., and Rosenbusch, J. P. (1986) *Methods Enzymol.* **125,** 309–328
32. Jokiranta, T. S., Tissari, J., Teleman, O., and Meri, S. (1995) *FEBS Lett.* **376,** 31–36
33. Vivien, D., and Wrana, J. L. (1995) *Exp. Cell Res.* **221,** 60–65
34. ten Dijke, P., Ichijo, H., Franzen, P., Schulz, P., Saras, J., Toyoshima, H., Heldin, C. H., and Miyazono, K. (1993) *Oncogene* **8,** 2879–2887
35. Sankar, S., Mahooti-Brooks, N., Bensen, L., McCarthy, T. L., Centrella, M., and Madri, J. A. (1996) *J. Clin. Invest.* **97,** 1436–1446
36. Centrella, M., Casinghino, S., Kim, J., Pham, T., Rosen, V., Wozney, J., and McCarthy, T. L. (1995) *Mol. Cell. Biol.* **15,** 3273–3281
37. Mulder, K. M., Segarini, P. R., Morris, S. L., Ziman, J. M., and Choi, H. G. (1993) *J. Cell. Physiol.* **154,** 162–174
38. McCaffrey, T. A., Consigli, S., Du, B., Falcone, D. J., Sanborn, T. A., Spokojny, A. M., and Bush, H. L., Jr. (1995) *J. Clin. Invest.* **96,** 2667–2675

Bicistronic Retrovirus Vectors Rapidly Generate Mammalian Cells Stably Expressing Multiple Genes at Predetermined Levels

Xuedong Liu,* Stefan N. Constantinescu,* Jonathan S. Bogan*#, David Hirsch* and Harvey F. Lodish*†§

* Whitehead Institute for Biomedical Research,
Nine Cambridge Center, Cambridge, MA 02142

Diabetes Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114

† Department of Biology, Massachusetts Institute of Technology,
Cambridge, MA 02139

§ corresponding author; mailing address:
Whitehead Institute for Biomedical Research
Nine Cambridge Center, Cambridge, MA, 02142
Phone (617) 258-5216
Fax (617) 258-6768
e-mail: lodish@wi.mit.edu.

Abbreviations footnote: TGF-ß: transforming growth factor beta; GFP, green fluorescent protein; FACS, fluorescent activated cell sorting; IRES, internal ribosome entry site,

ABSTRACT

Expression of cloned genes at desired levels in appropriate host cells is essential for protein production and for studying protein function. In practice, controlled levels of expression have been difficult to achieve, especially for cell lines with low transfection efficiency or when expression of multiple genes is required. We have developed a set of bicistronic retroviral vectors for stable expression of multiple cDNAs at predetermined levels; these contain an internal ribosomal entry site (IRES) followed by a cDNA encoding either green fluorescent protein (GFP) or a cell surface marker protein. The gene of interest is placed in a multiple cloning site 5' of the IRES sequence under the control of the retroviral long terminal repeat (LTR) promoter. We show that levels of expression of the genes upstream and downstream of the IRES are highly correlated, both in transiently transfected packaging cells and in stably infected target cells. Stably transfected populations of cells or clonal cell lines expressing specific amounts of a desired protein can be easily generated by fluorescent activated cell sorting (FACS) of cells expressing predetermined levels of the protein encoded by the second reading frame. Further, correlation of a biologic effect with the level of expression of the protein downstream of the IRES can provide strong evidence for the function of the protein placed upstream of the IRES. Conversely, selection for or against the function of the gene upstream of the GFP can be monitored by increased or decreased expression of the downstream protein. Finally, because there is no block to superinfection with multiple retroviral vectors, cell lines expressing several exogenous proteins can be identified and selected using combinations of GFP and cell surface markers placed downstream of the IRES.

INTRODUCTION

The ability to express cloned genes at controlled levels is important for investigating their biological function. Various systems have been developed to allow transient or stable expression of a gene of interest in cultured mammalian cells, yet it is very difficult to obtain expression of a desired protein at a predetermined level. The traditional approach of constructing a clonal cell line that stably expresses the gene of interest relies on recombination of the transfected DNA with the host genome. Unfortunately, in mammalian cells this process is inherently inefficient because the recombination frequency is quite low; furthermore the site of integration strongly affects the level of expression. Additionally, there is also no clear correlation between the degree of resistance to various drugs commonly employed for selection, such as geneticin (G418) or hygromycin, and the expression level of the exogenous protein. Finally, selection of clones that have integrated the exogenous DNA can take several weeks. Since one is often required to screen many stable clones to obtain a cell line with the desired level of expression, construction of stable cell lines by transfection followed by selection is often laborious and time consuming. Use of other selectable markers, such as dihydrofolate reductase and glutamine synthetase, has been used to drive amplification of exogenous genes, but in this case the amplified DNA is not stable and continuous presence of the drug

is required [1, 2].

Retroviral gene transfer offers an efficient alternative for stable expression of exogenous proteins. In particular, the development of packaging cell lines that produce high virus titers after transient transfection has allowed efficient and essentially quantitative infection of many different target cell lines

3-6

. Stable expression of a cloned cDNA promoted by the viral LTR occurs after integration of a retroviral vector into the host genome. Maximum levels of expression are typically present within two days after infection, at which point drug selection may be used. However, among individual cells within an infected population there is considerable variability in the amount of exogenous protein expressed, due in part to the use of many sites in the host genome for provirus integration. Additionally, multiple insertions of the provirus into the genome can occur, resulting in further cell-to-cell variability in the amount of protein expressed. Thus, the considerable advantages of the current ecotropic retroviral vectors are qualified by the wide variation (perhaps a hundred-fold, see below) in the level of expression of the cloned cDNA within individual cells.

We have developed a set of retroviral vectors that takes advantage of the spectrum of expression levels of cloned cDNAs while simultaneously maintaining the high efficiency of retroviral gene transfer. These vectors employ an encephalomyocarditis virus IRES followed by a quantitative selection marker, such as GFP or a cell surface marker, detectable, respectively, by its intrinsic fluorescence or by staining live cells with a fluorescent antibody. Because expression from the two reading frames in these bicistronic vectors is strongly correlated, we can use FACS sorting to obtain stably transfected populations of cells which express a desired protein at a predictable and predetermined level. Strong evidence for the function of the protein placed upstream of the IRES can be obtained by correlating the extent of a biological effect with the level of expression of the protein downstream of the IRES. Conversely, selection for or against functional expression of the gene upstream of the IRES can be monitored by increased or decreased expression of the downstream protein. Although our recombinant retroviruses express the ecotropic envelope glycoprotein, we can stably infect several non-murine cell lines by first making stable transfectants expressing the murine ecotropic receptor, a basic amino acid

transporter [7, 8]

Finally, because there is no block to superinfection with multiple packaged retroviral vectors, cell lines expressing several exogenous proteins can be identified and selected using combinations of GFP and cell surface markers placed downstream of the IRES. Thus, our system permits the rapid generation of clones or populations of cells that express one or more exogenous proteins at controlled levels.

RESULTS

Bicistronic retroviral vectors

While most eukaryotic mRNAs are monocistronic, initiation of translation from internal AUG start codons occurs in encephalomyocarditis virus (EMCV) and other picornaviruses. The 5' untranslated region of EMCV contains a cis-acting sequence element, termed an IRES, which permits initiation of translation at an internal AUG

sequence

Thus, RNA transcripts containing an IRES are polycistronic, and can be used to express several gene products from a single mRNA. We and others have used this strategy to

express multiple genes in the same cell 12-

15

As exemplified by pMX-IRES-GFP (Figure 1), our set of plasmid vectors contain the murine Moloney virus LTR and a packaging signal; the EMCV IRES is placed the between the polylinker/stuffer and a cDNA encoding a selectable marker protein. We employ three different selectable markers to allow quantitative measurement by flow cytometry: GFP, murine CD2 and human CD4. We utilize DNA segments encoding only the extracellular and transmembrane domains of these cell surface markers so as to avoid the possibility that their expression might alter the intracellular signaling events we want to study. Upon transfection of these plasmids into the ecotropic retroviral packaging cell line BOSC23 cells, high titers of recombinant retrovirus are produced and can be used to stably infect many murine cell lines.

To infect native non-murine cell lines, one must generate an amphotropic retrovirus. An alternative, which we prefer, is to introduce the cDNA encoding the

ecotropic receptor into the non-murine cells we wish to infect. This method has the advantage that it is safer to use ecotropic viruses than amphotropic viruses, especially when known oncogenes are being expressed. Also, this method permits us to use existing ecotropic packaging lines that are known to produce high titers of virus

4-6

. We have demonstrated that expression of the murine ecotropic receptor in non-murine cells greatly enhances their infectability, making them similar to murine cell lines. Table 1 lists several cell lines that can be infected at high efficiency provided that they stably express the murine ecotropic receptor.

Correlation between the levels of expression of genes upstream and downstream of the IRES

Although an IRES has been used to express multiple gene in the same cells

12, 13, 16-19

, a

quantitative relationship between expression levels of the genes upstream and downstream of the IRES has not been established. To address this issue, mouse CD2 was placed upstream of the IRES-GFP in the pMX-IRES-GFP vector, to construct the bicistronic pMX-mCD2-IRES-GFP retrovirus (Figure 2). This plasmid was transfected into the packaging BOSC23 cell line and virus supernatant was used to infect a mink lung cell line, L20, stably expressing the ecotropic receptor. The level of CD2 was determined by staining with a PE-labeled anti-CD2 antibody followed by FACS analysis; no staining is required for FACS analysis of GFP fluorescence. As shown in Figure 2, over a 50-fold range there is a remarkably good correlation between the levels of expression of GFP and CD2, both in transiently transfected BOSC23 cells (bottom panel) and stably transfected L20 cells (middle panel). No fluorescence from either GFP or CD2 is evident in control, non- transfected L20 cells (top panel). Therefore, if one selects high GFP expressing cells it is very likely that the same cells will also express high levels of mCD2. We obtained similar results using vectors in which murine CD2 or human CD4 is encoded by the downstream open reading frame, using GFP to monitor expression from the upstream reading frame (data not shown).

The micrographs in Figure 3 show that individual cells indeed exhibit a high correlation of expression of the genes placed upstream and downstream, respectively, of the IRES. L20 cells were infected with the bicistronic vector pMX-Smad3-C-Flag-IRES-GFP, encoding a Flag- tagged Smad3 protein and GFP. The 5% of the cells exhibiting the most GFP fluorescence were sorted and expanded in culture. As judged by immunofluorescence microscopy, all of these cells express the Smad3-C-Flag protein (Panel A). Further, in the same cells the level of expression of GFP is strongly correlated with that of the Smad3-C-Flag protein (Panel B).

The experiment in Figure 4 shows how a bicistronic vector can be used to isolate, within two days of infection, a population of infected Ba/F3 cells expressing high levels of the erythropoietin receptor. Growth of murine Ba/F3 cells is normally dependent on the presence of interleukin 3 (IL3); these cells can grow in the presence of erythropoietin (Epo) provided the Epo receptor (EpoR) is expressed at an appropriate level. To obtain Epo-dependent Ba/F3 cells, the erythropoietin receptor was cloned into pMX-IRES-GFP, forming the bicistronic retrovirus pMX-mEpoR-IRES-GFP; high titer retrovirus was used to infect the Ba/F3 cells. As determined by FACS analysis, parental cells exhibit no GFP fluorescence (Panel A), while two days after infection 21.16% of the cells were GFP positive (Panel B). One population of these infected cells was selected for Epo dependent growth, in the absence of IL3. As shown in Panel C, after selection these cells exhibit high GFP levels, roughly ten- fold greater than the average of the unsorted population of GFP- positive infected cells (Compare to Panel B). Thus, selection for the function of the gene cloned upstream of the IRES results in enrichment of cells that exhibit increased expression of the gene cloned downstream of the IRES.

Another population of these cells was flow sorted two days after infection in order to isolate those expressing the highest 0.1% level of GFP fluorescence intensity. This sorted population was then grown in IL3 for several weeks. Importantly, these cells continued to grow at the same rate, without a lag period, when placed in medium containing Epo but no IL3, indicating that they expressed sufficient EpoRs to support proliferation. These cells expressed the same high levels of GFP (Panel C) as those directly selected for Epo- dependent growth (Panel D). As judged by Western blotting, these cells expressed the same amount of the EpoR as did those selected for Epo- dependent proliferation (Panel E). Thus, cells selected two days after infection for expression of high levels of GFP - the gene placed downstream of the IRES - express high levels of the gene placed upstream of the IRES, EpoR in this example. In addition, this experiment further substantiates the excellent correlation between the expression level of the marker gene and that of the gene of interest.

Expression of multiple genes at predetermined levels using bicistronic vectors

Because there is no block to superinfection with multiple packaged retroviral vectors, and because our packaged retroviruses are replication defective, cell lines expressing predetermined levels of several exogenous proteins can be identified and selected using combinations of GFP and cell surface markers placed downstream of the IRES. Genes of interest can be cloned into pMX-IRES-GFP, pMX-IRES-CD2 or pMX-IRES-CD4. While the resulting viruses can be used to coinfect a cell, we found sequential infection to be more efficient. As illustrated in Figure 5, when cells were sequentially infected with pMX-IRES-GFP and pMX-IRES- CD2, about 1% of the cells expressed both markers at high levels. FACS sorting can be used to select populations expressing desired levels of both GFP and CD2 marker proteins; were these vectors to contain cDNAs upstream of the IRES then the encoded proteins, too, would be expressed at predetermined levels.

Analysis of gene function by infection with bicistronic vectors followed by quantitative cell sorting

Our bicistronic vectors can be used to isolate cells expressing different amounts of any gene of interest and thus to test the biological function of the encoded protein, as is illustrated by the experiment in Figure 6. We previously showed that adding a FLAG epitope tag to the carboxyl - terminus of the Smad 3 alters its function; the Smad3-C-Flag protein exhibits a dominant negative phenotype when overexpressed in mink lung epithelial cells and inhibits the antiproliferative effects of TGF-ß

14

. In the experiment in Figure 6, L20 cells were infected with a pMX--Smad3-C-Flag-IRES-GFP retrovirus. Two days after infection, six pools, each comprising five to thirty thousand cells, were sorted by FACS according to their GFP levels. These pools were expanded in culture and an equal number of cells from each pool were grown for two and half weeks in the presence of 100 pM of TGF-ß. Like parental cells, cells expressing low levels of GFP were killed by TGF-ß, while those expressing successively higher average levels of GFP were increasingly able to form colonies in the presence of TGF-ß. Thus the level of expression of Smad3-C-Flag, measured by the level of GFP expression, correlates with the ability of cells to grow in the presence of TGF-ß. In showing that the level of Smad3-C-FLAG determines the TGF-ß resistance phenotype, this experiment illustrates how these bicistronic vectors can be used to determine a biological function of an encoded protein.

DISCUSSION

We have combined highly efficient retroviral mediated gene transfer and quantitative selection markers in order to achieve expression of desired genes at predetermined and controlled levels. With this system, the gene of interest is cloned in the polylinker region of a retroviral vector upstream of the IRES sequence (Figure 1). Downstream of the IRES is a gene encoding a protein that can be quantified by FACS sorting, either GFP or the truncated cell surface proteins CD2 or CD4 missing their cytoplasmic domains. Plasmid DNA is transiently transfected into packaging BOSC or Phoenix cells; two days later the viral supernatant is collected and used to infect the target cells. Cells can be sorted for expression of the gene downstream of the IRES two days after infection, and thus it is possible to generate pools of cells expressing desired levels of expression of the downstream gene as soon as the cells can be grown in culture. In our experience pools of cells suffice for most experiments; it is usually unnecessary to isolate clones of infected cells. Indeed, since clonal variation may result in individual cell lines that do not recapitulate all phenotypes of the parental line, we prefer to work with pools of infected cells whenever possible.

Our most important result is the very high correlation, over a ~50 fold range, between the level of expression of the selectable gene, either GFP or CD2 or CD4, and the gene of interest. This occurs both after transient transfection of the packaging cell line and, importantly, in stably infected cells, and can most easily be demonstrated by FACS analysis or immunofluorescence microscopy of the levels of expression of the proteins encoded by the genes upstream and downstream of the IRES (Figures 2 and 3).

We provided several examples of how this system expedites functional analysis of cloned proteins. In one, we isolated a population of cells expressing high levels of the EpoR only two days after infection by an EpoR- encoding retrovirus by selecting cells expressing a high level of the downstream gene protein, GFP. The level of EpoR expression was the same as in a population of cells selected for growth in Epo. Conversely, selection for function of the EpoR gene, placed upstream of the IRES resulted in enrichment of cells that exhibited higher levels of expression of GFP, the gene cloned downstream of the IRES. In the second, we showed that increasing levels of expression of a mutant dominant negative signaling protein, Smad3-C-Flag, correlated with the ability of the cells to become resistant to the growth- inhibitory effects of TGF-ß. We did not need to measure the actual levels of Smad3-C-Flag expression, but rather measured the levels of the protein, GFP, downstream of the IRES in the pMX--Smad3-C-Flag-IRES-GFP retrovirus. In showing that the level of Smad3-C-FLAG determines the TGF-ß resistance phenotype, this experiment illustrates how these bicistronic vectors can be used to determine a biological function of an encoded protein.

Our experiments thus substantiate the excellent correlation between the expression level of the selection or marker gene and that of the gene of interest, as

measured by protein expression levels and by biologic function. This system offers several distinctive advantages over traditional expression systems.

First, it is easy to establish stable expression of the gene of interest in a variety of cell lines or cell types. For cell lines derived from human or other species that are not readily infectable with ecotropic murine retroviruses, one can first generate a stable line expressing the murine ecotropic receptor. This usually renders the cells readily infectable. Cells infected with retroviruses stably express the gene of interest within two days after infection, which can easily be detected; our unpublished experiments showed that this level of expression is maintained in culture at least for several weeks.

Second, using our retroviral expression system one can easily separate infected from non infected cells by FACS sorting, and one can obtain pools of cells expressing various levels of the gene of interest. Because viral titers are usually in the range of 5 x 10^5 to 10^6 infectious units per ml, one can obtain large number of cells in a matter of a few days. Compared to the use of drug selection to obtain stable pools of expressing cells, this method is simpler and more efficient especially since one does not need to find the optimal concentration of drug to use in the selection procedure.

Third, since the fluorescent marker used in the IRES-based retroviral vector is quantitative and highly correlated with the level of the expression of the gene upstream of the IRES, it is possible to sort the cells expressing the gene at predetermined and desirable levels and expand them for functional studies or industrial production of proteins of interest. We have used this system to generate stable lines of infected cells expressing the same amount of wild- type and point mutant Epo receptors (S. C., X.L., and H.F.L., unpublished data).

Fourth, this system provides a rapid way to generate stable expressing cell lines. High titers of virus can be harvested two days after transfection, and if used directly for infection the desired cells can be sorted after two more days. Thus, a total of only four days is required to isolate pools of cells expressing predetermined amounts of the desired protein.

Fifth, this system takes full the advantage of the random integration of retroviral DNAs into the cellular genome, which creates a wide spectrum in the level of expression of the introduced gene. By titrating the expression level of a particular gene of interest, one can build a quantitative relationship between the amount of protein expressed and a functional consequence. The advantage of the system is even more distinct when one wishes to express two proteins at predetermined levels in the same cells. By sequentially infecting or coinfecting with any combination of our bicistronic retroviruses pMX-IRES-GFP, pMX-IRES-CD2, or pMX-IRES-CD4, one can obtain cells expressing defined levels of GFP, CD2, or CD4, and thus presumably expressing two genes of interest at desired levels.

We have used this expression system in variety of functional studies in several experimental systems. The methodology proved to be very powerful and the level of protein expression was more than adequate for functional studies. Quantitative expression of the gene of interest proved particularly important in studies with dominant negative mutants, where the level of expression of the mutant protein influences the phenotype of the cells. The traditional approach of stable transfection, selection of many independent clones, and Western blotting to determine expression level of the encoded protein is quite laborious, and the particular clones one chooses to study can be subject to great variation. Quantitative selection of pools of cells expressing different levels of protein followed by functional assessment of a particular cellular response is less subjective.

Our system also allows functional selection for the gene cloned upstream of the IRES to be detected by a proportional increase in the expression of a fluorescent marker encoded by the gene downstream of the IRES. In other cases, stable expression of certain gene products can have a deleterious or toxic effect on cell proliferation; it is almost impossible to obtain stable lines expressing such genes by transfection and selection for expression of drug resistance. The system we describe allows one to monitor the level of GFP (or CD2 or CD4) to gain information about potential deleterious effects of expression of an upstream gene, seen as a decrease in expression of the fluorescent marker. One can use this system to obtain stable lines expressing a low level of such a toxic gene. Thus, such fluorescent markers can be used to pre-set the level of expression of a gene of interest, or to reveal a functional selection for or against expression of a gene of interest.

Finally, cDNA libraries made in the pMX-IRES-GFP vector have proven useful in functional expression cloning. The Chinese hamster ovary (CHO) cell mutant ldlB defines a gene required for multiple steps in the normal *medial* and *trans* Golgi associated processing of glycoconjugates; recently we used this retrovirus-based expression cloning system to clone a murine cDNA, *LDLB*, that corrects the pleiotropic mutant phenotypes of

In summary, we have developed a simple, efficient and quantitative system to allow expression of multiples genes in mammalian cells at predetermined and controlled levels. This system should have wide application in studying the function of genes or for industrial production of recombinant proteins.

MATERIALS AND METHODS

Cell lines

Mv1Lu mink lung epithelial cells CCL64 (ATCC) and its L20 derivative expressing the murine ecotropic receptor

21

were grown in minimal essential medium (MEM) supplemented with non-essential amino acids, penicillin, streptomycin, and 10% fetal bovine serum (FBS, GIBCO-BRL) as described

14

previously
The retroviral packaging cell line BOSC23

4

was
maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% FBS. Ba/F3 cells were grown in IMDM supplemented with WEHI conditioned medium

15

DNA Expression Constructs

The retroviral expression vector pMX

22 has been described previously. pMX-GFP was constructed by inserting the EcoRI-NotI fragment from pEGFP-N1 (Clontech) into EcoRI-NotI digested pMX. pMX-IRES-GFP1.0 was made by inserting the PvuII-NcoI fragment containing the IRES sequence from pCITE3a (Novagen) into EcoRI- blunted and NcoI- digested pMX-GFP. pMX-IRES-GFP1.0 was digested with EcoRI and NotI followed by treatment with the Klenow fragment of DNA polymerase to fill in the ends. This fragment containing the IRES-GFP was isolated and inserted into SalI- digested and end- filled pMX, resulting in the pMX-IRES-GFP vector. pMX-mCD2-IRES-GFP was constructed by inserting a PCR amplified mCD2 fragment into the BamHI digested and blunt- ended pMX-IRES-GFP1.0. The construction of pMX-Smad3-C-Flag-IRES-GFP1.0 has been described previously

14 . To construct pMX-IRES-CD2 and pMX-IRES-CD4, we first constructed a backbone pMX-IRES plasmid by amplifying the encephalomyocarditis virus internal ribosome entry site

11, 23 from the plasmid pIRES1neo (Clontech) using the 5' primer 5'-AAA CTC GAG CGG CCG CCA GCA CAG TGG CCA TGG CAT CTA GGG CGG CCA ATT CG-3' and the 3' primer 5'-TTT TGT CGA CTA CGT ACC GCG GGT TGT GGC AAG CTT ATC ATC G-3'. These primers were chosen so as to introduce NotI, BstXI, and NcoI sites at the 5' end of the PCR product, and SalI, SnaBI, and SacII sites at the 3' end of the PCR product. The PCR product was digested with NotI and SalI and cloned into these sites in the pMX retroviral vector to make pMX-IRES. Next, the extracellular and membrane-spanning domains of the murine CD2 cDNA

24 were amplified using the primers 5'-GAT CAG CTC ATG AAA TGT AAA TTC CTG GG-3' and 5'-CTT GCG GCC GCT TAA TTC TGC GCT GCG GC-3'. The resulting PCR product was digested with NotI, blunted using Klenow, and cloned into the SnaBI site of pMX-IRES to make pMX-IRES-CD2. To make pMX-IRES-CD4, the extracellular and membrane-spanning domains of the human CD4 cDNA

were amplified using the primers 5'-TTT CCG CGG CCC ACC ATG AAC CGG GGA GTC CCT TTT AGG C-3' and 5'-GTA GTC GAC TTA GCG CCT TCG GTG CCG GCA-3'. The PCR product was digested with SacII and SalI and cloned into these sites in pMX-IRES to make pMX-IRES-CD4. Construction of pMX-mEpoR-IRES-GFP has been described previously .

Transfection, Infection, FACS analysis, and Cell Sorting

The pMX--IRES-GFP vector DNA and various derivative constructs were transfected into BOSC23 cells as described previously

[4]. Infection of mink lung L20 cells and determination of the virus titer were described previously

[14]. To analyze the level of GFP and cell surface expression of mCD2, BOSC23 cells transfected with pMX-mCD2-IRES-GFP and L20 cells infected with virus were resuspended in PBS containing 10 mM EDTA and 10% FBS, followed by incubation with a 2 μg/ml PE-coupled monoclonal antibody against mCD2 (Pharmingen). After a 30 min incubation on ice, cells were washed three times with PBS containing 10 mM EDTA and 10% FBS and analyzed by FACScan (Becton-Dickinson). To isolate cells expressing different level of GFP, forty-eight hours after infection cells were scanned by a FACSVantage cell sorter (Becton-Dickinson) and subsequently sorted according to their GFP levels. Typically, five thousand cells at given average fluorescent level were combined into a pool and expanded for further study.

Immunofluorescence Mv1Lu cells stably expressing Smad3-C-Flag were seeded in wells of a 6-well plate with cover slips at the bottom. After forty-eight hours, the cells were fixed in 4% paraformaldehyde, permeabilized with 0.5% Triton X-100, and blocked with 200 μg/ml normal goat IgG (Jackson ImmunoResearch). Smad3 proteins were stained with 3 μg/ml M2 Flag antibody (Kodak) for 30 min at room temperature followed by 5 μg/ml Cy3- conjugated goat anti-mouse IgG (Jackson ImmunoResearch) for 30 min at room temperature. After extensive washing, the cover slips were mounted with Slowfade (Molecular Probes), and the fluorescent staining patterns were visualized with a Zeiss Axioskop fluorescence microscope.

ACKNOWLEDGMENTS

We are grateful to Xanxin Hua for helpful discussions and sharing reagents, Drs. Warren Pear, Toshio Kitamura, Robert Mckay and Claude Nicolau for reagents, Edmund Karam for technical help, and Glen Paradis and Mike Jennings for help with cell sorting. This work was supported in part by NIH grants CA-63260, HL 32262, HL41484, and DK47618 to H. F. L, by a grant to H.F.L. from Pfizer Corporation, and by Grant CDR 88-03014 from the National Science Foundation to the M.I.T. Biotechnology Process Engineering Center. X.L. was supported by a postdoctoral fellowship from the NIH; S.N.C. by a fellowship from the Medical Foundation; J.S.B. by NIH Physician Scientist Award DK02371; and D. H. by N.I.H. training grant 5T32 CA 09541.

FIGURE LEGENDS

Fig. 1. Schematic diagram of the pMX-IRES-GFP retroviral vector. The IRES sequence is indicated as shaded box with an arrow indicating the direction of translation. The segment containing the bacterial origin of replication and ampicillin resistant gene is indicated by a black box. The stippled box represents sequence encoding the green fluorescent protein; alternatively, in the pMX-IRES-CD2 and pMX-IRES-CD4 retroviral vectors the corresponding sequence encodes truncated CD2 and CD4 cell surface proteins. Open boxes with arrows indicate the viral LTR sequences. The open box represents a stuffer fragment containing multiple cloning sites.

Fig. 2. Tight correlation between levels of expression of genes placed upstream and downstream of the encephalomyocarditis virus internal ribosome entry site (IRES). (A) FACScan analysis of uninfected L20 cells stained with the PE-coupled anti-mCD2 antibody; GFP fluorescence is on the abscissa and CD2 fluorescence on the ordinate. (B) FACScan analysis of L20 cells two days after infection with a recombinant bicistronic virus expressing mCD2 and GFP, shown in panel D; staining is with the PE-coupled anti-mCD2 antibody. (C) FACScan analysis of BOSC cells transfected two days earlier with the same vector and stained with PE-coupled anti-mCD2 antibody. (D) Schematic diagram of the bicistronic retrovirus vector used in this experiment.

Fig. 3. Individual cells exhibit a high correlation of expression of the Smad3-C-Flag and GFP genes placed upstream and downstream, respectively, of the IRES. L20 cells were infected with the bicistronic vector pMX-Smad3-C-Flag-IRES-GFP and the 5% of the cells exhibiting the most GFP fluorescence were sorted and expanded in culture. A portion of the cells were fixed and then treated with the M2 Flag antibody followed by a Cy3- labeled secondary antibody, as detailed in Materials and Methods. The same field of cells was photographed using a rhodamine filter, to detect Cy-3- labeled Smad3-C-Flag protein (A) and a FITC filter to detect GFP.

Fig. 4. Selection of transfected Ba/F3 cells expressing high levels of the erythropoietin receptor (EpoR). Ba/F3 cells were infected with the bicistronic retrovirus pMX-mEpoR-IRES-GFP encoding both the mEpoR and GFP. As determined by FACS analysis parental cells exhibit no GFP fluorescence (Panel A) while two days after infection 21.16% of the cells were GFP positive (Panel B). Infected cells were divided into two populations. One was sorted for the top 0.1% GFP fluorescence and maintained in IMF medium plus WEHI-conditioned medium as a source of IL3. Cells in the second population were selected for those able to proliferate in Epo by culturing them in IMF medium containing 1 unit/ml Epo but no IL3; only cells expressing a functional mEpo receptor can grow in this medium. After two weeks FACS analysis was used to measure the level of GFP expression in these populations; as shown in Panels C and D, both populations expressed the same high level of GFP. The Western Blot analysis in Panel E shows the level of expression of the

EpoR in parental cells (A), those selected for Epo- dependent proliferation (B), and those selected for high GFP expression (C).

Fig. 5. Expression of two genes at predetermined levels using two bicistronic vectors. L20 cells were sequentially infected with pMX-IRES-GFP followed by pMX-IRES-CD2. Two days later the cells were stained with PE-coupled anti-mCD2 antibody; FACS analysis measured the levels of CD2 and GFP in individual cells.

Fig. 6. Quantitative analysis of expression of a dominant negative Smad3 protein on the antiproliferative effect of TGF-ß. (A) L20 cells were infected with pMX-Smad3-C-Flag - IRES-GFP and scanned for GFP levels two days after infection. Five to thirty thousand cells were divided into six pools according to their GFP levels. The levels of GFP expression of these pools, a - g, were measures after growth for 14 days; mean fluorescence values are indicated. (B) Five thousand cells from each pool were set in 10 cm tissue culture plates and incubated for three weeks in the presence of 50 pM TGF-ß. Colonies growing on the plates were fixed in 95% ethanol, stained with crystal violet for 1 hr., and photographed with a digital camera.

Table 1.

Non-murine cell lines stably expressing the murine ecotropic receptor and that can be infected with high efficiency by murine ecotropic retroviruses.

Species	Cell line	References
Human	HT1080-ER7	Reference 26
	HeLa-ETR	Unpublished observations
	Hep3B-ER hepatoma	Unpublished observations
	SKW 6.4 lymphoblastoid	Reference 21
	EJ Bladder Carcinoma	Reference 8
	U118MG Glioblastoma	Reference 27
Chinese Hamster	CHO-K1-ETR	Reference 20
Mink	Mv1Lu L20	Reference 14

1. Weidle, U.H., Buckel, P. & Wienberg, J. 1988. Amplified expression constructs for human tissue-type plasminogen activator in Chinese hamster ovary cells: instability in the absence of selective pressure. *Gene.* **66:** 193-203.
2. Bebbington, C.R., Renner, G., Thomson, S., King, D., Abrams, D. & Yarranton, G.T. 1992. High-level expression of a recombinant antibody from myeloma cells using a glutamine synthetase gene as an amplifiable selectable marker. *Biotechnology (N Y).* **10:** 169-75.
3. Landau, N.R. & Littman, D.R. 1992. Packaging system for rapid production of murine leukemia virus vectors with variable tropism. *J Virol.* **66:** 5110-3.
4. Pear, W.S., Nolan, G.P., Scott, M.L. & Baltimore, D. 1993. Production of high-titer helper-free retroviruses by transient transfection. *Proc Natl Acad Sci U S A.* **90:** 8392-6.
5. Soneoka, Y., Cannon, P.M., Ramsdale, E.E., Griffiths, J.C., Romano, G., Kingsman, S.M. & Kingsman, A.J. 1995. A transient three-plasmid expression system for the production of high titer retroviral vectors. *Nucleic Acids Res.* **23:** 628-33.
6. Naviaux, R.K., Costanzi, E., Haas, M. & Verma, I.M. 1996. The pCL vector system: rapid production of helper-free, high-titer, recombinant retroviruses. *J Virol.* **70:** 5701-5.
7. Kim, J.W., Closs, E.I., Albritton, L.M. & Cunningham, J.M. 1991. Transport of cationic amino acids by the mouse ecotropic retrovirus receptor. *Nature.* **352:** 725-8.
8. Albritton, L.M., Tseng, L., Scadden, D. & Cunningham, J.M. 1989. A putative murine ecotropic retrovirus receptor gene encodes a multiple membrane-spanning protein and confers susceptibility to virus infection. *Cell.* **57:** 659-66.
9. Paul, A.V., Yang, C.F., Jang, S.K., Kuhn, R.J., Tada, H., Nicklin, M., Krausslich, H.G., Lee, C.K. & Wimmer, E. 1987. Molecular events leading to poliovirus genome replication. *Cold Spring Harb Symp Quant Biol.* **52:** 343-52.

10. Jang, S.K., Krausslich, H.G., Nicklin, M.J., Duke, G.M., Palmenberg, A.C. & Wimmer, E. 1988. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J Virol.* **62:** 2636-43.
11. Jang, S.K., Davies, M.V., Kaufman, R.J. & Wimmer, E. 1989. Initiation of protein synthesis by internal entry of ribosomes into the 5' nontranslated region of encephalomyocarditis virus RNA in vivo. *J Virol.* **63:** 1651-60.
12. Sugimoto, Y., Aksentijevich, I., Gottesman, M.M. & Pastan, I. 1994. Efficient expression of drug-selectable genes in retroviral vectors under control of an internal ribosome entry site. *Biotechnology (N Y).* **12:** 694-8.
13. Gurtu, V., Yan, G. & Zhang, G. 1996. IRES bicistronic expression vectors for efficient creation of stable mammalian cell lines. *Biochem Biophys Res Commun.* **229:** 295-8.
14. Liu, X., Sun, Y., Constantinescu, S.N., Karam, E., Weinberg, R.A. & Lodish, H.F. 1997. Transforming growth factor beta-induced phosphorylation of Smad3 is required for growth inhibition and transcriptional induction in epithelial cells. *Proc Natl Acad Sci U S A.* **94:** 10669-74.
15. Constantinescu, S.N., Wu, H., Liu, X., Beyer, W., Fallon, A. & Lodish, H.F. 1998. The anemic Friend virus gp55 envelope protein induces erythroid differentiation in fetal liver colony-forming units-erythroid. *Blood.* **91:** 1163-72.
16. Li, X., Wang, W. & Lufkin, T. 1997. Dicistronic LacZ and alkaline phosphatase reporter constructs permit simultaneous histological analysis of expression from multiple transgenes. *Biotechniques.* **23:** 874-8, 880, 882.
17. Zhou, Y., Aran, J., Gottesman, M.M. & Pastan, I. 1998. Co-expression of human adenosine deaminase and multidrug resistance using a bicistronic retroviral vector. *Hum Gene Ther.* **9:** 287-93.
18. Morgan, R.A., Couture, L., Elroy-Stein, O., Ragheb, J., Moss, B. & Anderson, W.F. 1992. Retroviral vectors containing putative internal ribosome entry sites: development of a polycistronic gene transfer system and applications to human gene therapy. *Nucleic Acids Res.* **20:** 1293-9.
19. Ghattas, I.R., Sanes, J.R. & Majors, J.E. 1991. The encephalomyocarditis virus internal ribosome entry site allows efficient coexpression of two genes from a recombinant provirus in cultured cells and in embryos. *Mol Cell Biol.* **11:** 5848-59.
20. Chatterton, J.E., Hirsch, D., Schwartz, J.J., Bickel, P.E., Rosenberg, R.D. & Krieger, M. 1999. Expression cloning of LDLB, a gene essential for normal Golgi function and assembly of a novel protein complex. *Proc Natl Acad Sci U S A.* in press.
21. Baker, B.W., Boettiger, D., Spooncer, E. & Norton, J.D. 1992. Efficient retroviral-mediated gene transfer into human B lymphoblastoid cells expressing mouse ecotropic viral receptor. *Nucleic Acids Res.* **20:** 5234.
22. Onishi, M., Kinoshita, S., Morikawa, Y., Shibuya, A., Phillips, J., Lanier, L.L., Gorman, D.M., Nolan, G.P., Miyajima, A. & Kitamura, T. 1996. Applications of retrovirus-mediated expression cloning. *Exp Hematol.* **24:** 324-9.

23. Jackson, R.J., Howell, M.T. & Kaminski, A. 1990. The novel mechanism of initiation of picornavirus RNA translation. *Trends Biochem Sci.* **15:** 477-83.
24. Seed, B. & Aruffo, A. 1987. Molecular cloning of the CD2 antigen, the T-cell erythrocyte receptor, by a rapid immunoselection procedure. *Proc Natl Acad Sci U S A.* **84:** 3365-9.
25. Maddon, P.J., Littman, D.R., Godfrey, M., Maddon, D.E., Chess, L. & Axel, R. 1985. The isolation and nucleotide sequence of a cDNA encoding the T cell surface protein T4: a new member of the immunoglobulin gene family. *Cell.* **42:** 93-104.
26. Hua, X., Liu, X., Ansari, D.O. & Lodish, H.F. 1998. Synergistic cooperation of TFE3 and Smad proteins in TGF-beta induced transcription of the plasminogen activator inhibitor-1 gene. *Genes Dev.* **12:** 3084-3095.
27. Suzuki, H., Brown, G.D., Ohno, K. & Meruelo, D. 1996. Infection of human cells by murine ecotropic viruses: retroviral vectors carrying the hygromycin resistance-encoding gene. *Gene.* **170:** 255-9.

Activation of the Erythropoietin Receptor by the gp55-P Viral Envelope Protein Is Determined by A Single Amino Acid in Its Transmembrane Domain

Running title: EpoR activation through the transmembrane domain

Stefan N. Constantinescu*, Xuedong Liu*, Wendy Beyer*^, Amy Fallon*, Srinivasan Shekar&$, Yoav I. Henis¶*, Steven O. Smith&$, and Harvey F. Lodish*§%

* Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142

^ Department of Biology, Kenyon College, Gambier, OH 43022

§ Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

& Department of Molecular Biophysics and Biochemistry, Yale University, CT 06520-8114

$ Present address: Department of Biochemistry and Cell Biology, SUNY Stony Brook, NY 11794

¶ Department of Neurobiochemistry, The George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv 69978, Israel

% Corresponding author

Telephone 617-258-5216
Fax 617-258-6768
e-mail lodish@wi.mit.edu

The Spleen Focus Forming Virus (SFFV) gp55-P envelope glycoprotein specifically binds to and activates murine erythropoietin receptors (EpoRs) coexpressed in the same cell, triggering proliferation of erythroid progenitors and inducing erythroleukemia. Here we demonstrate specific interactions between the single transmembrane domains of the two proteins that are essential for receptor activation. The human EpoR is not activated by gp55-P but mutation of a single amino acid, L238 in its transmembrane sequence to its murine counterpart serine results in its ability to be activated. The converse mutation in the murine EpoR (S238L) abolishes activation by gp55-P. Computational searches of interactions between the membrane- spanning segments of murine EpoR and gp55-P provide a possible explanation: The face of the EpoR transmembrane domain containing S238 is predicted to interact specifically with gp55-P but not gp55-A, a variant which is much less effective in activating the murine EpoR. Mutational studies on gp55-P M390, which is predicted to interact with S238, provide additional support for this model. Mutation of M390 to isoleucine, the corresponding residue in gp55-A, abolishes activation, but the gp55-P M390L mutation is fully functional. gp55-P is thought to activate signaling by the EpoR by inducing receptor oligomerization through interactions involving specific transmembrane residues

Key words: Epo receptor signaling/ red cell formation/ SFFV- gp55 env proteins/ transmembrane coiled- coil interactions

Introduction

Activation of the erythropoietin receptor (EpoR)

(D'Andrea *et al.*,

1989) by erythropoietin (Epo) is required for survival, proliferation and maturation of erythroid progenitors into mature red cells

(Wu *et al.*,

1995) . The EpoR is activated by Epo- induced

dimerization (Figure 1) (Syed *et al.*, 1998; Watowich *et al.*, 1996), reminiscent of the activation of the growth hormone receptor by growth hormone

(De Vos *et al.*,

1992) . Bivalent monoclonal antibodies directed to the extracellular domain of the EpoR and small non-covalently dimerized peptides also dimerize the EpoR and induce proliferation of Epo- dependent cell lines and formation of burst forming unit erythroid (BFU-E) colonies

(Elliott *et al.*, 1996; Livnah *et al.*, 1996; Wrighton *et al.*, 1996) . A point mutation, R129C, in the extracellular domain of the EpoR renders the receptor constitutively active

(Yoshimura *et al.*,

1990b) due to formation of an intermolecular disulfide bond between membrane-proximal segments (Figure 1)

(Watowich *et al.*,

1994) . While activation of the EpoR by high levels

of Epo (i.e. Epo transgenic animals) (Semenza *et al.*,

1990) results in erythrocytosis, it does not induce emergence of leukemic cells. Expression of the constitutively active EpoR (R129C) in early blood progenitors results in Epo- independent red cell formation and leukemias of

different lineages (Longmore and Lodish, 1991;

Longmore *et al.*, 1992; Longmore *et al.*, 1993) . This suggests that constitutive activation of the EpoR leads to cell transformation.

The EpoR belongs to the type I superfamily of cytokine receptors. Upon dimerization it activates a number of intracellular signal transduction pathways which include JAK2, STAT5, protein tyrosine phosphatases SHP-1 and SHP-2, PI-3' kinase, MAP-kinase and

protein kinase C (Ihle *et al.*, 1995; Watowich *et al.*,

1996) . The particular signals required for proliferation, survival and erythroid differentiation functions are not understood as all of these pathways are activated by many cytokine receptors

(Socolovsky *et al.*,

1997) and growth factor receptor tyrosine kinases.

Both the polycythemic (P) and anemic (A) strains of the Spleen Focus Forming Virus (SFFV) induce erythroleukemia in adult mice

(Friend, 1957; MacDonald *et al.*, 1980; Mirand *et al.*, 1968) . SFFV-P also induces polycythemia (increased numbers of red cells), a condition reminiscent of human *Polycythemia vera,*

while SFFV-A does not (Tambourin *et al.*,

1979) . The only SFFV genes required for oncogenesis and red cell formation are those encoding the envelope proteins gp55-P and -

A (Linemeyer *et al.*, 1982; Wolff and Ruscetti, 1985) . gp55-P specifically binds to and activates

murine erythropoietin receptors (EpoRs) (Li *et al.*,

1990) , triggering abnormal proliferation of erythroid progenitors most of which undergo normal erythroid differentiation. This initial stage of Friend disease, Epo- independent polyclonal erythroblastosis

(Mirand *et al.*, 1968; Wolff and Ruscetti,

1985) , is followed by the emergence of malignant clones of abnormally proliferating erythroid progenitors due to subsequent Spi-1 overexpression and mutations in genes such as p53

(Moreau-Gachelin *et al.*,

1990) . Intracellular signals induced by constitutively active forms of EpoR cooperate with the Spi-1 transcription factor to inhibit erythroid differentiation and promote transformation to erythroleukemia

(Quang *et al.*,

1997)

Only gp55-P and not gp55-A can activate the murine EpoR to support proliferation of cell lines and to induce Epo- independent red cell formation from early erythroid progenitors. The human EpoR binds to but is not activated by gp55-P and cannot support Epo-

independent proliferation of cell lines (Hoatlin and Kabat, 1995; Showers *et al.*, 1993).

We previously showed that gp55-A can activate the EpoR, albeit quantitatively or qualitatively weaker than gp55-P (Constantinescu *et al.*, 1998). Expression of gp55-P in day 12.5 fetal liver erythroid progenitors resulted in Epo- independent formation of erythroid colonies from both BFU-E's and CFU-E's, while expression of gp55-A promoted only CFU-E's but not BFU-E's to proliferate and differentiate into erythroid colonies

(Constantinescu *et al.*,

1998) . The effect of both gp55-P and -A was mediated through the EpoR since EpoR$^{-/-}$ fetal liver progenitors were not stimulated by expression of gp55 proteins, but did form erythroid colonies when the EpoR was

expressed *via* retroviral transfection (Constantinescu *et al.*, 1998; Wu *et al.*, 1995). These data suggest that activation of the EpoR by gp55-A is different from that by gp55-P. Biochemical experiments have established that gp55 proteins do interact with the EpoR

(Li *et al.*, 1990; Zon *et al.*,

1992) . While ^{125}I- labeled Epo can become

crosslinked to the EpoR and to gp55-P (Ferro *et al.*, 1993) it cannot be crosslinked to gp55-A, suggesting that gp55-A and -P interact differently with the EpoR

(Tarr *et al.*,

1997)

gp55 proteins are recombinants between envelope proteins of xenotropic and ecotropic retroviruses (Amanuma *et al.*, 1983); an important feature is the absence of any

cytosolic domain (Amanuma *et al.*,

1989) with the carboxyl-terminus coinciding with the end of the transmembrane domain (reviewed in

(Kabat,

1989)). The differences between the biologic activities of gp55-P and gp55-A are due to nine sequence differences between gp55-P and gp55-A in the carboxyl- terminal region, including 5 amino acid differences and the addition of two leucine residues in gp55-P within the transmembrane domain

(Chung *et al.*,

1989) . Thus, differences between the transmembrane domains of gp55-P and -A were suggested to cause their different biologic

activities (Amanuma *et al.*, 1989; Chung *et al.*, 1989). While gp55-P is well expressed at the cell-

surface, gp55-A is not (Amanuma *et al.*, 1989;

Ruscetti *et al.*, 1981) . The two additional leucine residues in the transmembrane domain of gp55-P was suggested to be the sequence responsible for its efficient cell surface expression

(Amanuma *et al.*,

1989) but this di-leucine motif has not been examined in the absence of the other amino acid sequence differences between gp55-P and gp55-A or in the absence of other mutations. Differences in cell- surface expression can be important since only gp55 proteins that are expressed at the cell- surface are able to

stimulate the EpoR for mitogenesis (Ferro *et al.*,

1993; Li *et al.*, 1995) . On the other hand, gp55 mutants that had deletions in transmembrane residues but that were expressed well on the

cell surface have been isolated recently (Watanabe *et al.*, 1995) . A possible interaction between transmembrane domains of gp55-P and EpoR is further suggested by data showing that chimeric Epo receptors in which the membrane-spanning domain was replaced by that of the IL3 receptor were not activated by gp55-P but responded normally to Epo

(Zon *et al.*,

1992)

We hypothesized that the sequence of the transmembrane domain of gp55 proteins critically influences their biologic activity (polycythemia versus anemia) by interacting differently with the transmembrane domain of the EpoR. Such distinct interactions may translate into different receptor activation mechanisms. Also, we hypothesized that the human EpoR is not activated by gp55-P because of distinct sequences in its transmembrane domain. Here we show that the human EpoR can indeed become activated by gp55-P if one transmembrane residue (Leu 238) is mutated to Ser (as in the mouse counterpart) and, conversely, the mouse EpoR can no longer be activated by gp55-P if Ser 238 is mutated to

Leu (as in the human counterpart). Using molecular dynamics simulations coupled to mutagenesis we propose a model in which one face of the gp55-P transmembrane domain interacts with the EpoR transmembrane domain exposing Ser 238 in the interface. At one position of gp55 a methionine (in gp55-P) or leucine but not an isoleucine (in gp55-A) supports EpoR activation, and our calculations predict that this methionine or leucine interacts specifically with serine 238 of the EpoR. Because gp55-P proteins are disulfide-linked dimers on the cell surface (Gliniak *et al.*, 1991), by interacting with one gp55-P dimer two EpoRs can be brought in sufficient proximity to activate signaling.

Results

Mutations in the transmembrane domains of the murine and human Epo receptors do not modify the receptor's ability to transduce signals when activated by Epo

The human EpoR binds to but is not activated by gp55-P

(Hoatlin *et al.*, 1995; Zon *et al.*,

1992) . Overall the human and murine EpoRs are 82

% identical in sequence (Jones *et al.*, 1990) and differ in only three of the 21 amino acids in the membrane- spanning segment. We generated all possible human→ mouse and mouse→ human EpoR mutations in this region. To test the biologic activity of these mutants we utilized a bicistronic retroviral vector containing the Encephalomyocarditis Virus Internal Ribosome Entry Site (IRES) in which the mutant EpoR and green fluorescent protein (GFP) are translated from the same mRNA (Figure 2A)

(Liu *et al.*,

1997) . Vectors expressing the different EpoRs were transiently transfected into BOSC 23 packaging cells and the released retroviruses used to infect IL-3- dependent Ba/F3 cells. The population of cells selected for growth in Epo expressed high levels of GFP (Figure 2B) and high constant levels of EpoR protein (Figure 2C). Expression of all wild- type or mutant EpoRs tested supported normal growth in response to Epo addition (Figure 3A and B); typical examples are shown in Figure 3C. In other experiments infected cells, growing in IL-3, were sorted for the top 0.1% GFP fluorescence; all of these cells could grow in Epo and exhibited levels of the appropriate mutant EpoR and GFP similar to those in the cell populations depicted in

Figure 2B. All of the resultant pools of cells expressed the same numbers of EpoRs with the same affinity for Epo, as assayed by ^{125}I-Epo binding and Scatchard analysis (data not shown). Thus, none of the mutations affected receptor biogenesis or activation by Epo.

A Leu to Ser point mutation in the transmembrane sequence renders the human Epo receptor sensitive to activation by gp55-P

Cell populations expressing wild type or mutant EpoRs were then assayed for the ability to support gp55-P dependent proliferation in the absence of Epo; to this end cells were infected with retroviruses encoding gp55-P

(Constantinescu *et al.*,

1998) . Cells expressing gp55-P and the wild- type

human EpoR or mutants V236L and V239L were unable to proliferate in the absence of Epo. Importantly, cells expressing gp55-P and the human EpoR mutant L238S were able to grow in the absence of Epo; they did so at a rate similar to cells expressing gp55-P and the wild- type murine EpoR (Figure 3B and D).

A Ser to Leu point mutation in the transmembrane sequence renders the murine EpoR resistant to activation by gp55-P

With one exception, cells expressing all of the mutant murine EpoRs tested (Figure 3A) were able to grow in the absence of Epo when gp55-P was coexpressed. The exception was mutant murine EpoR S238L, which did not support activation by gp55-P. Thus, Ser 238 is crucial for the functional interaction between the EpoR and gp55-P. A hydroxyl group at position 238, however, is not essential: Cells coexpressing gp55-P and mutant murine EpoR S238A do grow in the absence of Epo. That a methyl side chain at position 238 suffices for activation by gp55 suggests that putative EpoR- gp55 interactions at this position involve stereochemical van der Waals packing and not the presence of a hydroxyl group on the EpoR. This is not surprising, since serine hydroxyl groups tend to hydrogen bond back to the i - 4 carbonyl in transmembrane helices, resulting in a van der Waals surface similar to that of alanine. Mutation of the other hydroxyl- containing amino- acids in the transmembrane domain of the murine EpoR (Thr 229, Ser 231 or Thr 242) to Ala or mutation of either threonine residue to valine did not affect activation of the EpoR by Epo or gp55-P (Figure 3A).

Only the mouse EpoR and the human EpoR L238S mutant are activated by gp55-P to support formation of erythroid colonies from CFU-E's in day 12.5 EpoR$^{-/-}$ fetal liver cells

In order to assess activation of the EpoR by gp55-P in primary erythroid progenitors we used fetal liver cells derived from EpoR$^{-/-}$ embryos. These animals die at day 12.5 due to

severe anemia but do contain erythroid progenitors that can be rescued by expression of

the EpoR (Wu *et al.*,

1995) . $EpoR^{-/-}$ progenitors allow us to assess the capacity of mutant EpoRs to respond to either Epo or gp55-P expression in the absence of endogenous wild- type EpoR. To this end we generated bicistronic retroviral vectors in which cDNAs encoding mutant EpoRs and gp55-P were coexpressed from the same mRNA (Figure 4A). BOSC 23 cells transfected with similar bicistronic vectors showed similar levels of expression of both proteins (data not shown

(Constantinescu *et al.*, 1998; Liu *et al.*,

1997)). Viral titers were measured by incubating five- fold dilutions of the packaged viruses with NIH 3T3 cells and then analyzing for the expression of the gp55 protein by Western blotting. This assay is appropriate since all viruses contain the same IRES- gp55 sequence. As shown in Figure 4D, the titers of the MuEpoR-IRES-gp55-P and MuEpoR S238L-IRES-gp55-P viruses, encoding murine EpoRs, were the same, but were 4 to 5- fold higher than the titers of the two viruses encoding the HuEpoR, HuEpoR-IRES-gp55-P and HuEpoR L238S-IRES-gp55-P. This may be due to differences between the 5' untranslated sequences upstream of the murine and human EpoR cDNAs.

Day 12.5 EpoR$^{-/-}$ fetal liver cells were infected by packaged retroviruses and subsequently plated in methylcellulose in the presence or absence of 3 U/ml Epo. Cells expressing gp55-P and either wild- type murine EpoR or MuEpoR S238L formed the same number of erythroid CFU-E colonies in the presence of Epo (Figure 4B columns 2 and 4), confirming that the S238L mutation did not affect surface expression of the receptor or its ability to respond to Epo stimulation. When plated in the absence of Epo, cells expressing the wild- type murine EpoR and gp55-P formed 5 to 6 fold more CFU-E colonies (Figure 4B, column 1) than did cells infected with MuEpoR-IRES-GFP virus (Figure 4B column 5) or with the control GFP virus, confirming the ability of gp55-P to activate the wild- type murine EpoR. Among the explanations for the observed lower numbers of CFU-E colonies induced by gp55-P versus Epo is that the gp55-P cDNA was cloned downstream of the IRES while the EpoR cDNA was cloned upstream of the IRES. Often the level of expression of genes cloned downstream of the IRES is lower than those cloned upstream. Second, the inherent intracellular signal induced by gp55-P is probably weaker than that induced by Epo, as also reflected by the smaller size of the colonies

induced by gp55-P versus Epo (Constantinescu *et al.*, 1998)

Importantly, only background levels of CFU-E colonies were formed by cells coexpressing gp55-P and the mutant EpoR S238L in the absence of Epo (Figure 4B, compare columns 3 and 1). The difference between the ability of MuEpoR and MuEpoR S238L to support gp55-P induced colony formation (112.93±17.23 CFU-E's per fetal liver, N=3; versus 52.7±11.86 CFU-E's per fetal liver, N=3) is statistically significant (Student *t* test, $p<0.005$). Importantly, when the infected cells were cultured in the absence of Epo, the MuEpoR-IRES-GFP virus induced only background level of colonies (column

5) showing that all colonies induced by MuEpoR-IRES-gp55-P virus were due to expression of gp55-P.

Similarly, cells expressing gp55-P and either wild- type human EpoR or HuEpoR L238S formed the same number of erythroid CFU-E colonies in the presence of Epo (Figure 4C columns 2 and 4). The lower absolute numbers of colonies induced by Epo in cells infected with HuEpoR coding viruses (Figure 4C, columns 2 and 4) versus cells infected with murine EpoR- coding viruses (Figure 4B, columns 2, 4, and 6) were expected, given the lower titers of HuEpoR- coding viruses (Figure 4D). However, only cells coexpressing gp55-P and the mutant HuEpoR L238S formed colonies in the absence of Epo (Figure 4C, column 3); those expressing the wild- type human EpoR did not (Figure 4C column 1). The difference between the ability of HuEpoR L238S and HuEpoR to support gp55-P induced colony formation (66.76±14.42 CFU-E's per fetal liver, N=5; versus 24.76 ± 4.67 CFU-E's per fetal liver, N=5) is statistically significant (Student *t* test, $p<0.005$).

Thus, both in cell lines and in primary erythroid cells the presence of a small amino acid - either serine or alanine - at one position in the center of the transmembrane sequence of the EpoR is crucial for efficient activation by gp55-P but not for activation by Epo.

Ser 238 of the EpoR and Met 390 of gp55-P may lie in the interface of interacting EpoR-gp55-P transmembrane helices

A computational search strategy (Adams *et al.*, 1995; Adams *et al.*, 1996; Surti *et al.*, 1998) was used to identify low energy 'clusters' of interacting EpoR- gp55-P transmembrane helices. The structures generated from this global search can be evaluated on the basis of the mutational and biochemical data. As described in the Materials and Methods section, the search calculates low energy structures using molecular dynamics and energy minimization protocols at different rotational orientations and interaxial spacings of transmembrane helices. The lowest energy structure (ΔG = -71 kcal/mole) that was found places S238 of the murine EpoR in the dimer interface interacting with M390 of gp55-P. This dimer structure is stabilized by interactions along the length of the interface (Figure

5A and D) and provides a direct explanation for the observed mutational results since S238 is critical for interaction of the murine EpoR with gp55-P and M390 is one of the six transmembrane residues which differ between gp55-P and gp55-A (Figure 6).

Importantly, no low energy clusters were obtained for interactions between the transmembrane domains of murine EpoR and gp55-A (not shown). A computational search on the gp55-P homodimer revealed one very low energy structure that is stabilized by interhelical interactions involving a different face of the gp55-P helix than that predicted to interact with the EpoR (Figure 5D and E). We hypothesize that each of the transmembrane domains in a cell surface gp55-P dimer (or higher oligomer) binds specifically to the transmembrane segment of one EpoR, dimerizing and activating the EpoR.

Mutagenesis of the gp55-P and gp55-A transmembrane domains reveals which of the differing transmembrane residues are important for EpoR activation

These computational searches suggest that the transmembrane domain of gp55-P has a stronger interaction with the EpoR transmembrane domain than does gp55-A. This would be a possible explanation for the poor ability of gp55-A to activate EpoR signaling and for the well- established results showing that the transmembrane sequence of gp55-P is crucial

for inducing polycythemia *in vivo* (Amanuma *et al.*,

1989; Chung *et al.*, 1989; Watanabe *et al.*, 1995)

While these molecular models are a provocative way to imagine a possible interaction they do not provide sufficient detail to suggest compensatory mutations in the EpoR and gp55. However, the biology of our system allowed us to explore further the precise sequences required for productive transmembrane interactions since sequence differences between gp55-P and gp55-A must be responsible for the different biologic effects of these proteins. The membrane- spanning segment of gp55-P differs from that of gp55-A at 6 positions, and also contains two leucine residues (L396 and 397) "deleted" in gp55-A (Figure 6A). To test our model, we generated all possible gp55-P → gp55-A mutations, as well as

gp55-P mutants missing one (Del L397) or both (Del L396, 397) of the "extra" leucine residues. As determined by the ability to induce Epo- independent growth of Ba/F3 cells expressing the murine EpoR, only gp55-P mutants M390I, Del L396, 397, Del 397, and L399I could not activate the EpoR (Figure 6B). Both M390 and L397 are on the face of the gp55-P transmembrane helix predicted to interact with the EpoR, forming a left handed coiled coil. However, two other amino acids, G391 and L399, are not predicted to be on this face of the transmembrane helix (Figures 5 D and E; see below).

The position of a methyl group at position 390 of the gp55-P transmembrane domain determines the polycythemic <u>versus</u> anemic phenotype.

As noted, an Ile at position 390 renders gp55-P unable to activate the EpoR. Ile has a β-branched side chain, lacks rotational freedom, and is known to have drastic effects on

packing of α helical coiled coils (Harbury *et al.*,

1993) . We therefore predicted that, unlike isoleucine, introduction of a Leu at position 390 (gp55-P M390L) would allow gp55-P to activate the murine EpoR. Indeed, the gp55-P M390L mutant was fully able to activate the murine EpoR (Figure 6B). The position of one methyl group, at carbon γ or at carbon β is all that differs between leucine and isoleucine. These results thus support the notion that the transmembrane segments of the two proteins interact in a highly specific manner. According to our model (Figure 5D), Met 390 of gp55-P binds specifically to the murine EpoR in the region of Ser 238. On the opposite face of the gp55-P transmembrane helix, a mutation of L399 to I abolishes activation. A residue common to both gp55-P and -A, G391, immediately following M390 in gp55-P, seems to be important for activation since the gp55-P G391L mutant cannot activate the EpoR (Figure 6B). However, these results can only be interpreted correctly if these mutant gp55 proteins are expressed similarly on the cell surface. This issue is critical, since gp55-A is poorly expressed on the cell- surface

(Amanuma *et al.*, 1989; Ruscetti *et al.*,

1981)

Cell surface expression of gp55 mutant proteins and their biologic activity.

Intracellular processing and cell- surface expression of gp55 proteins expressed in NIH 3T3 cells can be used as a metric for the intrinsic ability of these proteins to be

transported to the plasma membrane (Amanuma *et al.*, 1989; Ruscetti *et al.*, 1981; Watanabe *et al.*, 1990; Watanabe *et al.*, 1995) . In order to determine the capacity of the various gp55 mutants to be expressed at the cell surface we generated stably expressing NIH 3T3 cell lines by infection with high titer SFFV virus encoding gp55 proteins. The different cell lines were analyzed by Western blotting with goat anti-Rauscher gp55 antibodies and shown to express similar levels of gp55 proteins (not shown). Cells were then analyzed by immunofluorescence microscopy using a triple sandwich (monoclonal anti-gp55 antibody 7C10, biotinylated goat anti-rat IgG, and Cy3- streptavidin) as described in Material and Methods. All labeling steps were performed on live cells at 4°C

in order to exclusively label the cell surface. This generates a patchy labeling pattern, as the antibody- labeled proteins at the surface of live, unfixed cells are swept into micropatches by cross- linking via the secondary antibodies and Cy3- streptavidin, as previously described (Henis *et al.*, 1994)

As shown in Fig. 7 gp55-P was very well expressed at the cell- surface while gp55-A was not, thus confirming results of others (Amanuma *et al.*, 1989; Ruscetti *et al.*, 1981). Mutants gp55-P M390I and G391L were equally well expressed at the cell- surface but, importantly, both fail to activate the EpoR. Moreover, two other mutants which are also unable to activate the MuEpoR, gp55-P Del L396, L397 and gp55-P Del L397, are expressed significantly at the cell- surface, albeit at lower levels. Therefore, deleting both extra- leucines, L396, L397 in gp55-P or only L397 does not abolish cell- surface expression while it does abolish activation of the EpoR. In a previous study

(Amanuma *et al.*,

1989) a mutant gp55- like molecule which lacked the two extra- leucines was found to be absent on the cell-surface. However, in that case additional mutations reminiscent of the Friend murine leukemia virus envelope protein were also present in the transmembrane domain (i.e. S392P).

Interestingly the gp55-P L399I mutant failed to express at detectable levels on the cell-surface (Fig. 7) and also failed to activate the EpoR. L399 was predicted to be on the opposite face of the transmembrane helix from that involved in interaction with the EpoR (Figures 5D and E) and thus would not have been expected to influence activation of the EpoR. Its lack of cell- surface expression explains its inability to activate the EpoR (Fig. 6

B). These data clearly show that sequence specificity is required for activation of the EpoR by gp55 proteins as a number of gp55 mutants that cannot activate the EpoR do reach the cell surface. Mutation of G391 (the residue following M390) also abolishes activation; G391 is predicted to allow the gp55-P and EpoR transmembrane helices to cross in close contact. In neu* (activated form of the neu protein through a V->E transmembrane mutation) the Glycine residue (G665) following E664 is also critical for activation of neu, as glycine residues are found with high frequency in a number of transmembrane helices to allow close packing/ crossing between helices.

Discussion

Our key findings can best be summarized by a molecular model (Figure 1): each subunit of a dimeric plasma membrane gp55-P binds, through its membrane- spanning segment, a single EpoR polypeptide, dimerizing the receptor and thus activating EpoR signal transduction pathways. A more detailed molecular model for the specific interactions between the transmembrane domains of the EpoR and gp55-P proteins (Figure 5D) was derived from computational searches. The convergence of these unbiased searches on a low energy structure consistent with the experimental data was striking. First, Serine 238 of the murine EpoR is predicted to be at the interface with gp55-P; a leucine residue is at the same position in the human EpoR, which is not activated by gp55-P. Mutation of this one amino acid in the human EpoR to its murine counterpart [HuEpoR L238S] results in its ability to be activated by gp55-P. Conversely, mutating the serine in the murine EpoR to its human counterpart [MuEpoR S238L] abolishes activation by gp55-P. The computational searches of EpoR-gp55-P transmembrane interactions yielded a low energy structure in which S238 of the EpoR transmembrane domain interacts specifically with M390 of gp55-P. Support for this low energy structure was provided by mutational studies analyzing the differences between gp55-A and gp55-P. In gp55-A, which is much less effective than gp55-P in activating the murine EpoR, the residue at position 390 is isoleucine. We show that mutation of Met 390 in gp55-P to its counterpart in gp55-A (M390I) results in its inability to activate the murine EpoR. In contrast, replacing this methionine with a leucine allows full biological activity. Thus, the EpoR can be activated by a non- related protein through a novel mechanism which involves highly specific interactions between residues within the transmembrane domains.

Activation of the EpoR by dimerization

EpoR can be activated by Epo- induced dimerization

(Watowich *et al.*,

1994) , by dimerization induced by small mimetic

peptides (Wrighton *et al.*,

1996) , by bivalent monoclonal antibodies directed to the EpoR, or by a point mutation in the extracellular domain which constitutively

dimerizes the receptor (Watowich *et al.*, 1992; Yoshimura *et al.*, 1990b) . The three dimensional structure of the complex of two EpoR extracellular domains bound to Epo shows that receptor orientation a ligand is critical for the efficiency of signaling

(Syed *et al.*,

1998) . A different orientation of the extracellular domains of two EpoR monomers is induced by peptide mimetics

(Livnah *et al.*,

1996) . Both peptide- and Epo- induced EpoR dimer assemblies differ from that of the human growth hormone receptor bound by its ligand

(De Vos *et al.*, 1992; Livnah *et al.*, 1996; Syed *et al.*,

1998; Wrighton *et al.*, 1996) . It is likely that, in the R129C EpoR mutant, which is a disulfide linked dimer, the two EpoR extracellular domains will have a yet another orientation, as will those dimerized by bivalent anti-receptor antibodies. While the efficiency of signaling (the ratio between the amount of substance required for cell growth and the Kd for binding) may be dependent on receptor

orientation (Syed *et al.*,

1998) , these data show that more than one mode of receptor dimerization can induce signal transduction, cell proliferation, and red cell differentiation. Recently, the crystal structure of the extracellular domain of the EpoR in its unliganded form has unexpectedly also revealed a preformed dimer in which the individual membrane- spanning and intracellular domains would be too far apart to permit

transphosphorylation of JAK2 molecules (Livnah *et al.*, 1999) . An *in vivo* fragment complementation assay further supports a model in which ligand binding induces a conformational change that allows JAK2 to be brought into closer proximity and thus activated

(Remy *et al.*,

1999) . Assuming that this model were correct, our results would suggest that binding of a gp55-P dimer to the preformed EpoR inactive dimer is able to induce a conformational change or to further move the EpoR monomers in close proximity in such a way that JAK2 activation is possible. Indeed, we have shown (data not published) that gp55-P and gp55 mutants that induce Epo- independence also induce constitutive activation of JAK2.

Specific interactions between the EpoR and gp55-P transmembrane domains

The residues in the murine EpoR and gp55-P transmembrane domains that we identified experimentally as crucial for biological activity are the ones that are shown computationally to stabilize the lowest energy complex between the EpoR and gp55-P transmembrane domains. The productive binding involves one face of the gp55-P transmembrane domain which recruits one face of the transmembrane domain of the mouse EpoR (Figures 1 and 5D). The combined approach of computational searches and mutagenesis allows us to begin to see the details of this specific interaction.

The human EpoR is not activated by gp55-P

((Hoatlin *et al.*, 1995; Showers *et al.*,

1993) ; the murine and human EpoRs differ only in

three positions in the transmembrane domain and are 82% identical overall

(Jones *et al.*,

1990) . Strikingly, mutation of Leu 238 to Ser in the transmembrane domain rendered the human EpoR sensitive to activation by gp55-P. Conversely, mutation of Ser 238 of the murine EpoR to Leu abolished activation by gp55-P. Mutation of Ser 238 to Ala did not abolish activation, nor did mutation of other hydroxyl- containing amino-acid residues to Ala or Thr (Fig. 3 A and B). The key position seems to be in the middle of the transmembrane domain as mutation of I 227 or L 228 to Ser did not have any effect on activation by gp55-P or by Epo (data not shown).

Activation of the EpoR by gp55-P was also assayed in primary fetal liver erythroid progenitors isolated from EpoR$^{-/-}$ embryos. Epo receptors with mutations in their transmembrane domains were equivalent in their ability to respond to Epo and induce erythroid colony formation from CFU-E's. Only the murine EpoR and the mutant HuEpoR L238S could support CFU-E differentiation when activated by gp55-P. Like the human EpoR, the murine EpoR S238L mutant was not activated by gp55-P but responded normally to Epo. Therefore, both in cell lines and in primary erythroid progenitors the presence of a small amino acid - either serine or alanine - at one position in the center of the transmembrane sequence of the EpoR is crucial for efficient activation by gp55-P but not for activation by Epo.

The face of the gp55-P transmembrane domain (containing Met 390 and Leu 397) predicted to interact with the mouse EpoR is different from that predicted to form the gp55-P homodimer (Figure 5E). Sequence- specificity is further demonstrated by our results using gp55 mutant proteins in which residues that differ between the polycythemic and anemic strains were changed to the corresponding sequences of the anemic strain. Previously it has been suggested that gp55-A does not activate the EpoR because it is poorly expressed on the cell surface (Amanuma *et*

al., 1989; Ruscetti *et al.*, 1981) . Here we show (Fig. 7) that, indeed, gp55-A is not detected by our triple- sandwich immunofluorescence on the cell- surface while gp55-P is well expressed. Notably, several mutant gp55 proteins that do not activate the EpoR were expressed at high levels on the cell- surface (gp55-P M390I and gp55-P G391L) while others were expressed at lower but significant levels (gp55-P Del L396, L397 and gp55-P Del L397). The one mutation (L399I) that would not have been predicted to inhibit activation of the EpoR was not expressed on the surface, thus explaining its lack of activation of EpoR. These results are consistent with those reported

by (Watanabe *et al.*,

1995) which showed that deletion of three residues at a time from the transmembrane domain of gp55-P can abolish activity but without affecting cell- surface expression. Although the gp55-P mutant that lacks both extra-leucine residues (Del L396, L397) is expressed at slightly reduced levels on the cell surface we suggest that this is not the reason why this mutant is inactive. A previous study

(Amanuma *et al.*,

1989) reported the isolation of a gp55- like molecule which did not have the two extra-leucine residues and was not expressed significantly on the cell- surface. As this mutant gp55 contained additional amino acid changes we suggest that the two leucines are crucial for biologic activity and not for cell -surface expression. Insertion of the two extra- leucines into gp55-A does not rescue the cell- surface expression defect (Henis et al, personal communication). Taken together our data clearly show that stringent transmembrane sequence specificity is required for activation of the EpoR by gp55 proteins. More importantly, replacing the methionine at position 390 in gp55-P with leucine (M390L) does allow full activation of the EpoR, thus excluding a unique role for the sulfur atom of the methionine. Since an isoleucine at that position in gp55 (M390I) leads to a protein unable to activate the EpoR, the ability of gp55-P to activate the EpoR depends on whether the methyl group is attached to the β (Ile) or γ (Leu) carbon atom of the aliphatic amino acid at position 390. Ile is present at this position in gp55-A (Figure 6A), which does not activate the EpoR to promote Ba/F3 cell proliferation. Such specificity is consistent with the notion that Met 390 (or Leu 390) in gp55-P mediates activation through specific van der Waals packing interactions with EpoR in the region of Ser 238, possibly a left handed coiled coil arrangement of transmembrane helices.

Such subtle changes causing dramatic biologic effects are not without precedent. In other coiled coils, such as the GCN4 leucine zipper, replacement of Leu with Ile at the hydrophobic positions of the heptad repeat changes the oligomerization status of the protein from dimer to trimer or tetramer (Harbury *et al.*, 1993). Met and Leu were also shown to be compatible at a critical position in transmembrane helix 6 of rhodopsin, where Met is involved in a specific interaction with transmembrane helix 7 by which rhodopsin signaling is kept off in the dark. However, Met 257 cannot be changed to any residue other that Leu without a loss of basal 11-cis-retinal induced inhibition of signaling in the

dark (Han *et al.*,

1998)

Unique features of EpoR/ gp55-P transmembrane domain interactions

The interaction between the transmembrane domains of murine EpoR and gp55-P is translated into an array of biologic activities. On the structural level, we suggest that this interaction is novel since it involves a stereochemical fit between hydrophobic amino acids of two unrelated membrane proteins with single membrane- spanning helices. While there are a number of examples where charged and/ or polar residues are determinants of interactions between membrane domains, only in the case of glycophorin is the

homodimerization of two transmembrane domains generated entirely by van der Waals interactions.

The *neu* tyrosine kinase receptor can be activated by a Val to Glu point mutation in the transmembrane domain (Bargmann *et al.*, 1986) which dimerizes and constitutively activates

the receptor (Bargmann *et al.*, 1986; Burke *et al.*,

1997) . Achondroplasia, the most common genetic form of dwarfism, results from a Gly to Arg substitution in the transmembrane domain of the fibroblast growth factor receptor 3 (FGFR3)

(Rousseau *et al.*, 1994; Shiang *et al.*,

1994) causing constitutive activation of the receptor

(Webster and Donoghue,

1996) . Transmembrane helical interactions are crucial for the assembly of the T cell receptor (TCR); specific pairs of TCR chains (i.e. TCR α and CD-3δ) assemble due to interactions between specific charged sequences in

their transmembrane domains (Manolios *et al.*,

1990) . Complex formation between class II MHC α and β chains relies on residues with opposite charge at the amino-terminus of the membrane- spanning domains and on a structural motif involving several glycines on the

same face of the putative helices (Cosson and Bonifacino, 1992)

The ability of hydrophobic packing interactions to stabilize helix association has previously been seen in the glycophorin A transmembrane helix dimer

(Furthmayr and Marchesi, 1976; MacKenzie *et al.*,

1997) The glycophorin A transmembrane domain forms a stable homodimer through packing interactions involving a seven residue motif, LIxxGVxxGVxxT. Interestingly, the polar side chain of the threonine residue does not stabilize the dimer through interhelical hydrogen bonding, but rather the threonine contributes to dimerization through non- polar interactions involving the β- methyl group

(MacKenzie *et al.*,

1997) . As predicted in the EpoR for the β-hydroxyl group of Ser 238, the threonine β-hydroxyl group hydrogen bonds back to the i-4 carbonyl of the same helix. The interaction between the murine EpoR and gp55-P transmembrane domains is unique, however, since it involves a stereochemical fit between hydrophobic amino acids of two unrelated proteins.

Finally, many viruses encode small single- spanning membrane proteins that activate specific growth factor receptors and that may participate in cell transformation and/or

stimulate virus growth (Drummond-Barbosa and Di

Maio, 1997; Petti *et al.*, 1997) . The

fibropapillomavirus protein E5 forms a stable complex with and activates specifically the

PDGF β receptor (Petti *et al.*,

1997) . The determinants for specific activation of the PDGF receptor β are charged and polar sequences in the extracellular and transmembrane domains of the receptor. The interaction results in dimerization and

constitutive activation of the receptor (Petti *et al.*, 1997) . This interaction, like that between gp55-P and the EpoR, may prove to be a general mechanism by which unrelated protein oligomers function to dimerize cell- surface receptors via interactions within transmembrane α- helixes.

Materials and Methods

Generation of EpoR mutants

The murine and human EpoR cDNAs were cloned in the pMX-IRES-GFP bicistronic retroviral vector upstream of the IRES

(Constantinescu *et al.*, 1998; Liu *et al.*,

1997) . The translation of the two proteins is tightly linked in that expression of GFP is proportional over a 100- fold range to the level of expression of the protein encoded by the cDNA placed upstream of the IRES

(Liu *et al.*,

1997) . All mutant Epo receptor constructs were generated by PCR overlap extension using as external primers vector- and IRES- derived sequences. Because the human EpoR receptor has, compared to the murine, one extra amino acid in its extracellular domain, the numbering of HuEpoR residues was adjusted to that of the murine for easy comparison. After sequencing, plasmid DNAs were used to transfect retroviral packaging cells in order to generate high titer retroviruses. gp55-P cDNA, kindly provided by Dr. Sandra Ruscetti (National Cancer Institute, Frederick, MD) was used as a template to generate gp55-P to -A mutants by PCR overlap extension.

Generation of retroviral supernatants

High titer replication free retroviral supernatants were generated by transient transfection of the BOSC packaging cell line (Pear *et al.*, 1993). Of each retroviral construct, 5 μg of DNA was transfected by the calcium phosphate method

(Constantinescu *et al.*,

1998) . After collection of viral supernatants, expression in BOSC cells of various EpoR or gp55 mutants was measured by Western blots

as described (Constantinescu *et al.*,

1998) , while expression of GFP was measured by fluorescence activated cell sorting (FACS). Viral titers were measured by FACS as

described (Constantinescu *et al.*,

1998)

Assay for Epo- dependent and gp55-P dependent proliferation

IL3- dependent Ba/F3 cells growing in RPMI medium supplemented with 10% fetal bovine medium, antibiotics, and 5% WEHI supernatant (as a source of IL3) were infected with bicistronic viruses encoding GFP and different EpoRs (murine, human, wild type or mutants) for 4 h in the presence of 4 μg/ml polybrene. Infected cells were scanned by FACS for GFP fluorescence; typically infection efficiencies were 9-15%. For every infected construct, cells were either selected directly in Epo (0.5-1 unit/ ml) for

proliferation by removing IL3 or, alternatively, cells were first sorted for the top 0.1 % GFP fluorescence and then placed in Epo in the absence of IL3. The capacity to support proliferation in Epo was measured for each transduced EpoR construct; both selected and sorted cells gave similar results for each construct. All experiments were performed on pools of cells and not on clones. The expression of the EpoR was confirmed by Western blot analysis using antibodies directed to the murine EpoR (C-187,

(Yoshimura *et al.*,

1990a) and human EpoR (C-20, Santa Cruz Biotechnology, Inc., Santa Cruz, CA). Cells expressing different EpoRs were infected with

viruses encoding gp55-P as described (Constantinescu *et al.*, 1998) and the ability of transduced gp55-P to induce Epo- independent proliferation of cells was assayed by removing Epo or WEHI medium and placing cells in 24 well plates in medium without growth factor, as described

(Constantinescu *et al.*,

1998) . Epo- independent cells were expanded and tested for gp55-P expression; high levels of gp55-P were demonstrated in cells growing in the absence of Epo by Western blot analysis using an anti Rauscher-gp70 goat antibody (National Cancer Institute serum repository) recognizing the xenotropic- derived region of gp55 proteins or using a monoclonal anti-gp55 antibody (7C10, gift of Dr. Sandra Ruscetti; data not shown).

Erythroid colony formation in primary fetal liver cells

Day 12.5 homozygous EpoR$^{-/-}$ fetal liver cells were harvested and single cell suspensions were prepared as described (Constantinescu *et al.*, 1998; Wu *et al.*, 1995). Cells were infected with various bicistronic retroviruses in the presence of 4 μg/ml polybrene for 4 hours in Iscove's modified Dulbecco's medium (IMDM) containing 20% fetal bovine serum. For Colony Forming Unit- Erythroid (CFU-E) assays, 10^5 nucleated progenitors were plated in semisolid 0.1% methylcellulose medium (MethoCult 3230, Stem Cell Technologies, Vancouver, British Columbia, Canada) containing 20% fetal bovine serum, as described

(Constantinescu *et al.*, 1998; Wu *et al.*, 1995) in the presence or absence of 3 units/ml Epo. Colonies generated by CFU-E progenitors were scored 72 hours after plating by staining with diaminobenzidine (Sigma, St. Louis, MO). The efficiency of infection for a particular set of bicistronic retroviruses was represented by the number of colonies induced by Epo since EpoR$^{-/-}$progenitors absolutely require the EpoR to differentiate into red cells

(Wu *et al.*,

1995) . As previously described

(Constantinescu *et al.*,

1998) when viruses encoding both EpoR and GFP were used to infect parallel aliquots of fetal liver progenitors which were cultured in suspension in the presence of Epo and scanned by FACS for GFP expression 36 h after infection; 10-15%, infection efficiencies were typically obtained.

Computational searches

The computational search strategy has been described previously by Brünger and colleagues (Adams *et al.*, 1995; Adams *et al.*, 1996; Lemmon *et al.*, 1994; Nilges and Brunger, 1993; Treutlein *et al.*, 1992). The method is applicable to either homo- or heterodimers of interacting helices. Canonical α-helices were generated from the transmembrane sequences of gp55-P (residues 382-404), gp55-A (residues 382-404) and the EpoR (residues 226-249). The calculations on gp55-A were carried with and without the insert of Leu 396 and Leu 397, present in the gp55-P sequence. Heterodimer searches were carried on the EPOR-gp55-P and EPOR-gp55-A helices, while homodimer searches were carried out on the gp55-P and gp55-A helices. Low energy conformations were identified by rotating each helix in the helix dimer through rotation angles φ1 and φ2

from 0 - 360°. At 45° increments, molecular dynamics simulations were performed using simulated annealing of all atomic coordinates. The structures were energy minimized before and after the molecular dynamics simulations. The parameters used for the molecular dynamics simulations and energy minimization were the same as those used by

Adams et al. (Adams *et al.*,

1995) . The starting geometries included both left-handed (+25°) and right-handed (-50°) crossing angles. The crossing angles, as well as the rotation angles φ1 and φ2, were allowed to vary during the simulation. This allows the helices to adopt low-energy coiled-coil geometries. However, in individual simulations the distance between the two helix axes was held fixed at 10.0, 10.5 or 11.0 Å. There was no translational offset between the helices. Four different molecular dynamics simulations

were carried out for each starting geometry. During the cycle of molecular dynamics and minimization, the helix dimers migrate from their initial geometries and can group together to form well-defined clusters of structures. A cluster is defined as a group of at least ten structures where the root mean square deviation of the atom positions is less than 1 Å between any given structure in the cluster and the next most similar structure. For each cluster, the individual dimer structures are averaged and then energy minimized to yield a single low energy dimer structure. This analysis leads to a small pool of possible structures that can then be evaluated on the basis of experimental data.

Immunofluorescence microscopy

NIH 3T3 cells stably expressing different gp55 mutants were washed twice with Dulbecco's PBS supplemented with 1% BSA (Sigma) and incubated in the same buffer (45 min, 4°C) with goat γ– globulin (200 μg/ml) to block nonspecific staining. This was followed by successive incubations (4°C, 1 h each, with three washes between incubations) with: (a) 7C10 rat monoclonal antibody (ascites) (1:200 dilution); (b) biotinylated goat anti- rat IgG (20 μg/ml); and (c) Cy3- streptavidin (2 μg/ml). The 7C10 antibodies were a generous gift from Dr. Sandra K. Ruscetti, National Cancer Institute, Frederick, MD, while the biotinylated goat anti-rat IgG and Cy3- streptavidin were from Jackson ImmunoResearch Laboratories (West Grove, PA). All incubations with antibodies were carried out on live cells in the cold; under these conditions the antibodies cannot penetrate the cells and only cell-surface proteins exposed to the extracellular medium can be labeled. The cells were then washed and fixed successively in methanol (-20°C, 5 min) and acetone (-20°C, 2 min), and mounted with Prolong antifade mounting solution (Molecular Probes, Eugene, OR). Fluorescence digital images were acquired with a Nikon Eclipse 800 microscope (X100 oil immersion objective) coupled to a CCD camera (Micromax, Princeton Instruments), using OpenLab (by ImproVision) software. All images were taken under identical conditions, exported in TIFF format to Adobe Photoshop, and printed.

ACKNOWLEDGMENTS

We thank Drs. Stephanie Watowich and Rebecca Wells for reagents and discussions; Dr. Merav Socolovsky for critical comments on the manuscript; Dr. Sandra K. Ruscetti, National Cancer Institute, Frederick, MD for the generous gift of SFFV plasmids and monoclonal 7C10 antibodies; Stream Wang for excellent technical assistance; and Glenn Paradis, MIT/CCR Central Flow Cytometry Laboratory, for invaluable help with FACS sorting and analysis. We thank Dr. Merton Bernfield and Olga Goldberg, Children's Hospital, Harvard Medical School, Boston, MA, for their generous gift of reagents. This research is supported by Grant HL 32262 from The National Institutes of Health and by a grant from Amgen Corporation to HFL. S.N.C. held a fellowship from the Anna Fuller Fund and is now a fellow of the Medical Foundation/ Charles A. King Trust. X.L. holds a postdoctoral fellowship from the National Institutes of Health.

References

Adams, P. D., Arkin, I. T., Engelman, D. M., and Brunger, A. T. (1995): Computational searching and mutagenesis suggest a structure for the pentameric transmembrane domain of phospholamban. *Nat Struct Biol* **2**, 154-62.

Adams, P. D., Engelman, D. M., and Brunger, A. T. (1996): Improved prediction for the structure of the dimeric transmembrane domain of glycophorin A obtained through global searching [published erratum appears in Proteins 1997 Jan; 27(1):132]. *Proteins* **26**, 257-61.

Amanuma, H., Katori, A., Obata, M., Sagata, N., and Ikawa, Y. (1983): Complete nucleotide sequence of the gene for the specific glycoprotein (gp55) of Friend spleen focus-forming virus. *Proc Natl Acad Sci U S A* **80**, 3913-7.

Amanuma, H., Watanabe, N., Nishi, M., and Ikawa, Y. (1989): Requirement of the single base insertion at the 3' end of the env-related gene of Friend spleen focus-forming virus for pathogenic activity and its effect on localization of the glycoprotein product (gp55). *J Virol* **63**, 4824-33.

Bargmann, C. I., Hung, M. C., and Weinberg, R. A. (1986): Multiple independent activations of the neu oncogene by a point mutation altering the transmembrane domain of p185. *Cell* **45**, 649-57.

Burke, C. L., Lemmon, M. A., Coren, B. A., Engelman, D. M., and Stern, D. F. (1997): Dimerization of the p185neu transmembrane domain is necessary but not sufficient for transformation. *Oncogene* **14**, 687-96.

Chung, S. W., Wolff, L., and Ruscetti, S. K. (1989): Transmembrane domain of the envelope gene of a polycythemia-inducing retrovirus determines erythropoietin-independent growth. *Proc Natl Acad Sci U S A* **86**, 7957-60.

Constantinescu, S. N., Wu, H., Liu, X., Beyer, W., Fallon, A., and Lodish, H. F. (1998): The anemic Friend virus gp55 envelope protein induces erythroid differentiation in fetal liver colony-forming units-erythroid. *Blood* **91**, 1163-72.

Cosson, P., and Bonifacino, J. S. (1992): Role of transmembrane domain interactions in the assembly of class II MHC molecules. *Science* **258**, 659-62.

D'Andrea, A. D., Lodish, H. F., and Wong, G. G. (1989): Expression cloning of the murine erythropoietin receptor. *Cell* **57**, 277-85.

De Vos, A. M., Ultsch, M., and Kossiakoff, A. A. (1992): Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science* **255**, 306-311.

Drummond-Barbosa, D., and Di Maio, D. (1997): Virocrine transformation. *Biochim Biophys Acta* **1332**, M1-17.

Elliott, S., Lorenzini, T., Yanagihara, D., Chang, D., and Elliott, G. (1996): Activation of the erythropoietin (EPO) receptor by bivalent anti-EPO receptor antibodies. *J Biol Chem* **271**, 24691-7.

Ferro, F., Jr., Kozak, S. L., Hoatlin, M. E., and Kabat, D. (1993): Cell surface site for mitogenic interaction of erythropoietin receptors with the membrane glycoprotein encoded by Friend erythroleukemia virus. *J Biol Chem* **268**, 5741-7.

Friend, C. (1957): Cell-free transmission in adult Swiss mice of a disease having the character of a leukemia. *J. Exp. Med.* **105**, 307-318.

Furthmayr, H., and Marchesi, V. T. (1976): Subunit structure of human erythrocyte glycophorin A. *Biochemistry* **15**, 1137-44.

Gliniak, B. C., Kozak, S. L., Jones, R. T., and Kabat, D. (1991): Disulfide bonding controls the processing of retroviral envelope glycoproteins. *J Biol Chem* **266**, 22991-7.

Han, M., Smith, S. O., and Sakmar, T. P. (1998): Constitutive activation of opsin by mutation of methionine 257 on transmembrane helix 6. *Biochemistry* **37**, 8253-61.

Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993): A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **262**, 1401-7.

Henis, Y. I., Moustakas, A., Lin, H. Y., and Lodish, H. F. (1994): The Type II and Type III TGF-ß receptors form homo-oligomers. *J. Cell Biol* **126**, 139-154.

Hoatlin, M. E., Ferro, F., Jr., Geib, R. W., Fox, M. T., Kozak, S. L., and Kabat, D. (1995): Deletions in one domain of the Friend virus-encoded membrane glycoprotein overcome host range restrictions for erythroleukemia. *J Virol* **69**, 856-63.

Hoatlin, M. E., and Kabat, D. (1995): Host-range control of a retroviral disease: Friend erythroleukemia. *Trends Microbiol* **3**, 51-7.

Ihle, J. N., Witthuhn, B. A., Quelle, F. W., Yamamoto, K., and Silvennoinen, O. (1995): Signaling through the hematopoietic cytokine receptors. *Annu Rev Immunol* **13**, 369-98.

Jones, S. S., D'Andrea, A. D., Haines, L. L., and Wong, G. G. (1990): Human erythropoietin receptor: cloning, expression, and biologic characterization. *Blood* **76**, 31-5.

Kabat, D. (1989): Molecular biology of Friend viral erythroleukemia. *Curr Top Microbiol Immunol* **148**, 1-42.

Lemmon, M. A., Treutlein, H. R., Adams, P. D., Brunger, A. T., and Engelman, D. M. (1994): A dimerization motif for transmembrane alpha-helices. *Nat Struct Biol* **1**, 157-63.

Li, J.-P., D'Andrea, A. D., Lodish, H. F., and Baltimore, D. (1990): Activation of cell growth by binding of Friend spleen focus-forming virus gp55 glycoprotein to the erythropoietin receptor. *Nature* **343**, 762-4.

Li, J. P., Hu, H. O., Niu, Q. T., and Fang, C. (1995): Cell surface activation of the erythropoietin receptor by Friend spleen focus-forming virus gp55. *J Virol* **69**, 1714-19.

Linemeyer, D. L., Menke, J. G., Ruscetti, S. K., Evans, L. H., and Scolnick, E. M. (1982): Envelope gene sequences which encode the gp52 protein of spleen focus-forming virus are required for the induction of erythroid cell proliferation. *J Virol* **43**, 223-33.

Liu, X., Sun, Y., Constantinescu, S. N., Karam, E., Weinberg, R. A., and Lodish, H. F. (1997): Transforming growth factor beta-induced phosphorylation of Smad3 is required for growth inhibition and transcriptional induction in epithelial cells. *Proc Natl Acad Sci USA* **94**, 10669-10674.

Livnah, O., Stura, E. A., Johnson, D. L., Middleton, S. A., Mulcahy, L. S., Wrighton, N. C., Dower, W. J., Jolliffe, L. K., and Wilson, I. A. (1996): Functional mimicry of a protein hormone by a peptide agonist: the EPO receptor complex at 2.8 A. *Science* **273**, 464-71.

Livnah, O., Stura, E. A., Middleton, S. A., Johnson, D. L., Jolliffe, L. K., and Wilson, I. A. (1999): Crystallographic evidence for preformed dimers of erythropoietin receptor before ligand activation. *Science* **283**, 987-990.

Longmore, G. D., and Lodish, H. F. (1991): An activating mutation in the murine erythropoietin receptor induces erythroleukemia in mice: a cytokine receptor superfamily oncogene. *Cell* **67**, 1089-102.

Longmore, G. D., Pharr, P., and Lodish, H. F. (1992): Mutation in murine erythropoietin receptor induces erythropoietin-independent erythroid proliferation in vitro, polycythemia in vivo. *Leukemia* **6**, 130S-134S.

Longmore, G. D., Pharr, P., Neumann, D., and Lodish, H. F. (1993): Both megakaryocytopoiesis and erythropoiesis are induced in mice infected with a retrovirus expressing an oncogenic erythropoietin receptor. *Blood* **82**, 2386-95.

MacDonald, M. E., Reynolds, F. H., Jr., Van de Ven, W. J., Stephenson, J. R., Mak, T. W., and Bernstein, A. (1980): Anemia- and polycythemia-inducing isolates of Friend spleen focus-forming virus. Biological and molecular evidence for two distinct viral genomes. *J Exp Med* **151**, 1477-92.

MacKenzie, K. R., Prestegard, J. H., and Engelman, D. M. (1997): A transmembrane helix dimer: structure and implications. *Science* **276**, 131-3.

Manolios, N., Bonifacino, J. S., and Klausner, R. D. (1990): Transmembrane helical interactions and the assembly of the T cell receptor complex. *Science* **249**, 274-7.

Mirand, E. A., Steeves, R. A., Lange, R. D., and Grace, J. T., Jr. (1968): Virus-induced polycythemia in mice: erythropoiesis without erythropoietin. *Proc Soc Exp Biol Med* **128**, 844-9.

Moreau-Gachelin, F., Ray, D., de Both, N. J., van der Feltz, M. J., Tambourin, P., and Tavitian, A. (1990): Spi-1 oncogene activation in Rauscher and Friend murine virus-induced acute erythroleukemias. *Leukemia* **4**, 20-3.

Nilges, M., and Brunger, A. T. (1993): Successful prediction of the coiled coil geometry of the GCN4 leucine zipper domain by simulated annealing: comparison to the X-ray structure. *Proteins* **15**, 133-46.

Pear, W. S., Nolan, G. P., Scott, M. L., and Baltimore, D. (1993): Production of high-titer helper-free retroviruses by transient transfection. *Proc Natl Acad Sci U S A* **90**, 8392-6.

Petti, L. M., Reddy, V., Smith, S. O., and Di Maio, D. (1997): Identification of amino acids in the transmembrane and juxtamembrane domains of the platelet-derived growth factor receptor required for productive interaction with the bovine papillomavirus E5 protein. *J Virol* **71**, 7318-27.

Quang, C. T., Wessely, O., Pironin, M., Beug, H., and Ghysdael, J. (1997): Cooperation of Spi-1/PU.1 with an activated erythropoietin receptor inhibits apoptosis and Epo-dependent differentiation in primary erythroblasts and induces their Kit ligand-dependent proliferation. *Embo J* **16**, 5639-53.

Remy, I., Wilson, I. A., and Michnick, S. W. (1999): Erythropoietin receptor activation by a ligand-induced conformation change. *Science* **283**, 990-993.

Rousseau, F., Bonaventure, J., Legeai-Mallet, L., Pelet, A., Rozet, J. M., Maroteaux, P., Le Merrer, M., and Munnich, A. (1994): Mutations in the gene encoding fibroblast growth factor receptor-3 in achondroplasia. *Nature* **371**, 252-4.

Ruscetti, S. K., Feild, J. A., and Scolnick, E. M. (1981): Polycythaemia- and anaemia-inducing strains of spleen focus-forming virus differ in post-translational processing of envelope-related glycoproteins. *Nature* **294**, 663-5.

Semenza, G. L., Dureza, R. C., Traystman, M. D., Gearhart, J. D., and Antonarakis, S. E. (1990): Human erythropoietin gene expression in transgenic mice: multiple transcription initiation sites and cis-acting regulatory elements. *Mol Cell Biol* **10**, 930-8.

Shiang, R., Thompson, L. M., Zhu, Y. Z., Church, D. M., Fielder, T. J., Bocian, M., Winokur, S. T., and Wasmuth, J. J. (1994): Mutations in the transmembrane domain of FGFR3 cause the most common genetic form of dwarfism, achondroplasia. *Cell* **78**, 335-42.

Showers, M. O., De Martino, J. C., Saito, Y., and D'Andrea, A. D. (1993): Fusion of the erythropoietin receptor and the Friend spleen focus-forming virus gp55 glycoprotein transforms a factor-dependent hematopoietic cell line. *Mol Cell Biol* **13**, 739-48.

Socolovsky, M., Dusanter-Fourt, I., and Lodish, H. F. (1997): The prolactin receptor and severely truncated erythropoietin receptors support differentiation of erythroid progenitors. *J Biol Chem* **272**, 14009-14012.

Surti, T., Klein, O., Aschheim, K., DiMaio, D., and Smith, S. O. (1998): Structural models of the bovine papillomavirus E5 protein. *Proteins: Structure, Function and Genetics* **33**, 601-612.

Syed, R. S., Reid, S. W., Li, C., Cheetham, J. C., Aoki, K. H., Liu, B., Zhan, H., Osslund, T. D., Chirino, A. J., Zhang, J., Finer-Moore, J., Elliott, S., Sitney, K., Katz, B. A., Matthews, D. J., Wendoloski, J. J., Egrie, J., and Stroud, R. M. (1998): Efficiency of signalling through cytokine receptors depends critically on receptor orientation. *Nature* **395**, 511-516.

Tambourin, P. E., Wendling, F., Jasmin, C., and Smadja-Joffe, F. (1979): The physiopathology of Friend leukemia. *Leuk Res* **3**, 117-29.

Tarr, K., Watowich, S. S., and Longmore, G. D. (1997): Cell surface oligomerization of the erythropoietin receptor complex differs depending on its mode of activation. *J Biol Chem* **272**, 9099-9107.

Treutlein, H. R., Lemmon, M. A., Engelman, D. M., and Brunger, A. T. (1992): The glycophorin A transmembrane domain dimer: sequence-specific propensity for a right-handed supercoil of helices. *Biochemistry* **31**, 12726-32.

Watanabe, N., Nishi, M., Ikawa, Y., and Amanuma, H. (1990): A deletion in the Friend spleen focus-forming virus env gene is necessary for its product (gp55) to be leukemogenic [published erratum appears in J Virol 1990 Nov; 64(11):5694]. *J Virol* **64**, 2678-86.

Watanabe, N., Yugawa, T., Ikawa, Y., and Amanuma, H. (1995): Both the changes of six amino acids and the C-terminal truncation caused by a one-base insertion in the defective env gene of Friend spleen focus-forming virus significantly affect the pathogenic activity of the encoded leukemogenic membrane glycoprotein (gp55). *J Virol* **69**, 7606-11.

Watowich, S. S., Hilton, D. J., and Lodish, H. F. (1994): Activation and inhibition of erythropoietin receptor function: role of receptor dimerization. *Mol. Cell. Biol.* **14**, 3535-3549.

Watowich, S. S., Wu, H., Socolovsky, M., Klingmuller, U., Constantinescu, S. N., and Lodish, H. F. (1996): Cytokine receptor signal transduction and the control of hematopoietic cell development. *Ann Rev Cell Dev Biol* **12**, 91-128.

Watowich, S. S., Yoshimura, A., Longmore, G. D., Hilton, D. J., Yoshimura, Y., and Lodish, H. F. (1992): Homodimerization and constitutive activation of the erythropoietin receptor. *Proc Natl Acad Sci U S A* **89**, 2140-4.

Webster, M. K., and Donoghue, D. J. (1996): Constitutive activation of fibroblast growth factor receptor 3 by the transmembrane domain point mutation found in achondroplasia. *Embo J* **15**, 520-7.

Wolff, L., and Ruscetti, S. (1985): Malignant transformation of erythroid cells in vivo by introduction of a nonreplicating retrovirus vector. *Science* **228**, 1549-52.

Wrighton, N. C., Farrell, F. X., Chang, R., Kashyap, A. K., Barbone, F. P., Mulcahy, L. S., Johnson, D. L., Barrett, R. W., Jolliffe, L. K., and Dower, W. J. (1996): Small peptides as potent mimetics of the protein hormone erythropoietin. *Science* **273**, 458-64.

Wu, H., Liu, X., Jaenisch, R., and Lodish, H. F. (1995): Generation of committed erythroid BFU-E and CFU-E progenitors does not require erythropoietin or the erythropoietin receptor. *Cell* **83**, 59-67.

Yoshimura, A., D'Andrea, A. D., and Lodish, H. F. (1990a): The friend spleen focus-forming virus glycoprotein gp55 interacts with the erythropoietin receptor in the

endoplasmic reticulum and affects receptor metabolism. *Proc. Natl. Acad. Sci. USA* **87**, 4139-4143.

Yoshimura, A., Longmore, G., and Lodish, H. F. (1990b): Point mutation in the exoplasmic domain of the erythropoietin receptor resulting in hormone-independent activation and tumorigenicity. *Nature* **348**, 647-9.

Zon, L. I., Moreau, J. F., Koo, J. W., Mathey-Prevot, B., and D'Andrea, A. D. (1992): The erythropoietin receptor transmembrane region is necessary for activation by the Friend spleen focus-forming virus gp55 glycoprotein. *Mol Cell Biol* **12**, 2949-

57.

FIGURE LEGENDS

Fig. 1. Mechanisms of activation of the EpoR. The natural ligand, Epo, dimerizes the receptor and induces red cell formation. The receptor can be covalently dimerized and constitutively activated by the R129C point mutation in the extracellular domain

(Yoshimura *et al.*,

1990b) , which results in Epo- independent formation of red cells and in leukemias of multiple lineages. Coexpression in the same cell of the gp55-P viral envelope protein activates the EpoR

(Li *et al.*,

1990) and is oncogenic only for the erythroid lineage since it requires expression of the EpoR. Cell surface gp55-P is a dimer

(Gliniak *et al.*,

1991) ; the present work demonstrates specific binding to the EpoR, mediated by the membrane- spanning domains, that induces EpoR oligomerization and Epo- independent activation.

Fig. 2. Expression of EpoR mutants in Ba/F3 cells.

(**A**) Schematic showing the bicistronic pMX-EpoR-IRES-GFP vectors encoding mutant EpoRs and GFP.

(**B**) Ba/F3 cells growing in IL3 containing medium were infected with bicistronic retroviruses encoding the wild- type human EpoR (HuEpoR) or the indicated mutants. The

efficiency of infection averaged 9-15%, as judged by the fraction of cells positive for GFP. Cells were washed and cultured in 0.5- 1 units/ml Epo; over 95% of cells growing for 2-3 days in Epo were positive for GFP. Similar results were obtained when cells were first grown in IL-3 and then sorted for the top 0.1% GFP fluorescence; these cells could grow in Epo and exhibited levels of the appropriate mutant EpoR and GFP similar to those depicted in this panel.

(**C**) Western blot analysis using antibodies directed against the human EpoR showing that pools of Ba/F3 cells expressing various mutant human EpoRs and selected for growth in Epo (panel b) contain similar numbers of Epo receptors.

Fig. 3. Activation by Epo and by gp55-P of EpoR mutants in Ba/F3 cells. Mutations introduced in the murine (**A**) and human EpoR (**B**) transmembrane domains and the effect of these mutations on the ability of receptors to support proliferation in the presence of Epo or upon coexpression of gp55-P and culture in the absence of Epo. Because the human EpoR receptor has, compared to the murine, one extra amino acid in its extracellular domain, the numbering of HuEpoR residues was adjusted to that of the murine for easy comparison.

(**C**) Indistinguishable Epo sensitivity for proliferation of Ba/F3 cells expressing the wild type human EpoR (HuEpoR), HuEpoR mutant L238S, and murine EpoR (MuEpoR). Ba/F3 cells growing in 1% WEHI (as a source of IL3) were infected with IRES-GFP bicistronic viruses encoding HuEpoR, HuEpoR L238S or MuEpoR. After 3 days the cells were switched from medium containing WEHI to that with 1 u/ml Epo, as detailed in the legend to Fig. 2 and in Material and Methods. After 48 h growth the cells were washed three times in RPMI and then cultivated in RPMI supplemented with 10% fetal bovine serum and the indicated concentrations of Epo (0.001- 10 u/ml). Cells expressing HuEpoR (rectangles), HuEpoR L238S (diamonds) and MuEpoR (solid circles) were counted at 72 h

after addition of Epo (at the indicated concentrations) on a Coulter Cell Counter. Each point is the average of four replicates ± standard deviation.

(**D**) Epo- independent growth of pools of Ba/F3 cells coexpressing MuEpoR or HuEpoR L238S with gp55-P. As shown in panels (**A**) and (**B**), coexpression of gp55-P with MuEpoR or HuEpoR L238S results in Epo- independence. Two independent pools of cells coexpressing MuEpoR and gp55-P (solid circles and open triangles) or two independent pools of cells coexpressing HuEpoR L238S and gp55-P (open triangles and diamonds) were incubated at a density of 75,000/ml in RPMI supplemented with 10% bovine calf serum but no other growth factor (*i.e.* IL3 or Epo) and counted on a Coulter Cell Counter at the indicated time points.

Fig. 4. gp55-P expression activates the wild- type murine EpoR and a mutant HuEpoR L238S to promote erythroid colony formation from day 12.5 EpoR$^{-/-}$ fetal liver CFU-E's.

(**A**) Schematic showing the bicistronic vectors encoding mutant EpoRs and gp55-P.

(**B**) The number of CFU-E colonies generated in the absence of Epo (-Epo; hatched columns) or in the presence of 3 units/ml Epo/ml (+ Epo; solid columns) in day 12.5 EpoR$^{-/-}$ fetal liver CFU-E's after infection by viruses encoding the wild type or S238L mutant murine Epo receptors (MuEpoR) and gp55-P, wild type MuEpoR and GFP in place of gp55-P, or only GFP (control). Fetal liver cells were infected with different viruses having similar titers (as assayed by Western blot analysis for gp55 expression on infected NIH 3T3 cells, panel D) for 4 h at 37 °C. Cells were plated in methylcellulose and CFU-E's were scored by staining with benzidine 72 h after plating. Results are expressed as colonies induced per fetal liver in the presence or absence of Epo. Data represent the mean of 4 assays (columns 1 - 4 and 6) or 3 assays (columns 5 and 7) ± 1 standard

deviation. All EpoR- encoding viruses induced the same number of CFU-E colonies when the infected cells were cultured in the presence of Epo (columns 2, 4, and 6). In contrast, the number of CFU-E colonies induced by the MuEpoR- IRES- gp55-P virus when the infected cells were cultured in the absence of Epo (column 1) was significantly greater (Student *t* test, $p<0.005$) than those induced by the MuEpoR S238L - IRES- gp55-P virus (column 3) or by the MuEpoR- IRES- GFP virus which expresses GFP in place of gp55-P (column 5).

(**C**) The number of CFU-E colonies formed in the absence of Epo (-Epo; hatched columns) or in the presence of 3 units/ml Epo/ml (+ Epo; solid columns) in day 12.5 $EpoR^{-/-}$ fetal liver CFU-E's after infection by viruses encoding the wild type or L238S mutant human Epo receptors (HuEpoR) and gp55-P, or only GFP (control). Infection and detection of CFU-E's colonies were as in Panel B; results are expressed as colonies induced per fetal liver in the presence or absence of Epo. Data represent the mean of 5 independent assays ± 1 standard deviation. Both EpoR- encoding viruses induced the same number of CFU-E colonies when the infected cells were cultured in the presence of Epo (columns 2 and 4). In contrast, the number of CFU-E colonies induced by the HuEpoR L238S - IRES- gp55-P virus when the infected cells were cultured in the absence of Epo (column 3) was significantly greater (Student *t* test, $p<0.005$) than those induced by the wild- type HuEpoR - IRES- gp55-P virus (column 1). The latter figure was not significantly different from the number of CFU-E colonies induced by the control GFP virus (column 5).

(**D**) Expression of gp55 protein in NIH 3T3 cells infected with virus encoding MuEpoR-IRES-gp55-P, MuEpoR S238L-IRES-gp55-P, HuEpoR-IRES-gp55-P and HuEpoR-L238S-IRES-gp55-P. Typical titration experiments were performed by incubating five fold dilutions of packaged virus with NIH 3T3 cells (50,000 cells/ well) for 8 h in the presence of 8 μg/ml polybrene. Cells were counted after 72 h and analyzed by Western blotting

with an anti-gp55 antibody. Shown is a typical example of an infection with undiluted virus.

Fig. 5. Modeling of transmembrane interactions between the murine EpoR and gp55-P proteins using molecular dynamics simulations and energy minimization.

(**A**) EpoR - gp55-P interactions. Energy contributions of EpoR residues 226 to 249 calculated for the complex between the transmembrane domains of the EpoR and gp55-P proteins. The average structure shown in panel D is the lowest energy complex of eight clusters generated in the conformational search. The interaxial distance was fixed at 10.5 Å.

(**B**) EpoR - gp55-P interactions. Energy contributions of gp55-P residues 382 - 404 calculated for the complex shown in panel D between the transmembrane domains of the EpoR and gp55-P proteins. Comparison of panels (A) and (B) with the structure shown in panel D indicates that the low energy contribution of Thr 383 results from hydrogen-bonding interactions with Ser 231 of EpoR. In contrast, Ser 238 of EpoR hydrogen-bonds back to Leu 234 rather than across the dimer interface. The energy contributions of Ser 238 and Met 390 (on gp55-P) results predominantly from van der Waals interactions.

(**C**) gp55-P - gp55-P interactions. Energy contributions of gp55-P residues 382 - 404 calculated for the homodimer complex shown in panel E. The average structure shown in panel E is the lowest energy complex of ten clusters generated in the conformational search. The interaxial distance was fixed at 10.5 Å. Conformational searches with interaxial separations of 9.5 and 10.0 Å led to fewer clusters of structures. The lowest energy clusters had the orientation of gp55-P helices seen in panel E and an increased packing contribution for Gly 391.

(**D**) Details of the predicted interaction between the transmembrane domains of the murine EpoR and gp55-P; Met 390 of gp55-P is predicted to interact at the interface with Ser 238 and Leu 234 of the murine EpoR.

(**E**) Predicted structure of the homodimer of the gp55-P transmembrane domain; Gly 391 is placed in the interface of the gp55-P homodimer. The face of the gp55-P helix predicted to interact with the murine EpoR is opposite to that predicted to form the homodimer.

Fig. 6. Identification of the precise transmembrane sequences of gp55-P which differ from gp55-A that are involved in activation of the EpoR signaling.

(**A**) Sequence alignment between the transmembrane sequences of gp55-P and gp55-A.

(**B**) The indicated gp55-P mutants were tested in proliferation assays for their capacity to induce Epo- independent growth in murine EpoR- expressing Ba/F3 cells. Except for M390L and G391L all mutations convert residues in gp55-P to those in gp55-A.

Fig. 7. Cell-surface expression of gp55 mutant proteins.

NIH 3T3 cells stably expressing various gp55 mutants at similar levels (as judged by Western blot using anti-gp55 goat serum, not shown) were processed for immunofluorescence microscopy as described under Material and Methods. All labeling steps were performed on live cells at 4° C, in order to label exclusively cell-surface proteins. All photographs were taken in the same session, under identical conditions and exposure times, and the film rolls were developed together and printed under identical conditions to enable direct comparison. Bar = 10 μm. For each cell line the percentage of fluorescent- labeled cells was evaluated by counting 500 cells in several fields and scoring cells visibly labeled above the background level. Samples with 20-30% labeled cells were

marked "++", those with 10-15% labeled cells were designated "+"; and samples with 2% or fewer labeled cells were marked "-".